Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 120 (2010) 1748–1775

stochastic processes and their applications

www.elsevier.com/locate/spa

On Malliavin's differentiability of BSDEs with time delayed generators driven by Brownian motions and Poisson random measures

Łukasz Delong[a,*], Peter Imkeller[b]

[a] *Institute of Econometrics, Division of Probabilistic Methods, Warsaw School of Economics, Al. Niepodleglosci 162, 02-554 Warsaw, Poland*

[b] *Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany*

Received 20 July 2009; received in revised form 15 April 2010; accepted 2 May 2010
Available online 10 May 2010

Abstract

We investigate solutions of backward stochastic differential equations (BSDEs) with time delayed generators driven by Brownian motions and Poisson random measures, that constitute the two components of a Lévy process. In these new types of equations, the generator can depend on the past values of a solution, by feeding them back into the dynamics with a time lag. For such time delayed BSDEs, we prove the existence and uniqueness of solutions provided we restrict on a sufficiently small time horizon or the generator possesses a sufficiently small Lipschitz constant. We study differentiability in the variational or Malliavin sense and derive equations that are satisfied by the Malliavin gradient processes. On the chosen stochastic basis this addresses smoothness both with respect to the continuous part of our Lévy process in terms of the classical Malliavin derivative for Hilbert space valued random variables, as well as with respect to the pure jump component for which it takes the form of an increment quotient operator related to the Picard difference operator.

Keywords: Backward stochastic differential equation; Time delayed generator; Poisson random measure; Malliavin's calculus; Canonical Lévy space; Picard difference operator

* Corresponding author. Tel.: +48 22 5648617; fax: +48 22 5648617.
E-mail addresses: lukasz.delong@sgh.waw.pl (Ł. Delong), imkeller@mathematik.hu-berlin.de (P. Imkeller).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2010.05.001

1. Introduction

Introduced in [20], backward stochastic differential equations have been thoroughly studied in the literature during the last decade; see [12] or [14] and references therein. Viewed from the perspective of Peng who interprets their key structural feature as a *nonlinear conditional expectation*, the close link to the stochastic calculus of variations or Malliavin's calculus becomes apparent. In fact, in a Clark–Ocone type formula, the control component of the solution pair of a BSDE with a classical globally Lipschitz generator without time delay on a Gaussian basis turns out to be the Malliavin trace of the other component; see Proposition 5.3 in [12] or Theorem 3.3.1 in [14]. Not only this observation attributes an important role to Malliavin's calculus in the context of stochastic control theory and BSDE. As the simplest example, let us recall that hedging strategies in complete market models corresponds to Malliavin derivatives of wealth processes; see [16]. The fine structure and sensitivity properties of solutions of BSDEs or systems of forward and backward stochastic differential equations have been approached by means of the stochastic calculus of variations (see [2,1]), and applied to provide explicit descriptions of delta hedges of insurance related financial derivatives in [3]. Let us mention that Malliavin's calculus has been applied to prove regularity of trajectories and thus to provide a first numerical scheme for BSDEs with generators of quadratic growth; see for instance [15]. More generally, it has been established as a key tool in the numerics of control theory and mathematical finance, for instance to enhance the convergence speed of discretization schemes for solutions of BSDEs, see [17,18]. BSDEs have proved to be an efficient and powerful tool in a variety of applications in stochastic control and mathematical finance. In all of these applications, variational smoothness of their solutions is fundamental for describing their properties.

In this spirit, and with the aim of clarifying smoothness in the sense of the stochastic calculus of variations and related properties of BSDEs in a more general setting, in this paper we study the equations with dynamics given for $t \in [0, T]$ by

$$Y(t) = \xi + \int_t^T f(s, Y_s, Z_s, U_s)\mathrm{d}s - \int_t^T Z(s)\mathrm{d}W(s) - \int_t^T U(s,z)\tilde{M}(\mathrm{d}s, \mathrm{d}z).$$

An equation of this type will be called BSDE with time delayed generator. It is driven by a Lévy process, the components of which are given by a Brownian motion and a Poisson random measure. In these new types of equations, a generator f at time s depends in some measurable way on the past values of a solution $(Y_s, Z_s, U_s) = (Y(s+u), Z(s+u), U(s+u, .))_{-T \le u \le 0}$. Very recently, time delayed BSDEs driven by Brownian motion and with Lipschitz continuous generators have been investigated for the the first time in [8], and in more depth in [10]. We would like to refer the interested reader to the accompanying paper [10], where existence and uniqueness questions are treated, and examples given in which multiple solutions or no solutions at all exist. Further, several solution properties are investigated, including the comparison principle, measure solutions, the inheritance property of boundedness from terminal condition to solution, as well as the *BMO* martingale property for the control component. We would like to point out that all results from [10] can be extended and proved in the setting of this paper.

Our main findings are the following. First, we prove that a unique solution exists, provided that the Lipschitz constant of the generator is sufficiently small, or the equation is considered on a sufficiently small time horizon. This is the extension of Theorem 2.1 from [10] to be expected. Secondly, we establish Malliavin's differentiability of the solution of a time delayed BSDE, both with respect to the continuous component of the Lévy process, which coincides with the classical Malliavin derivative for Hilbert-valued random variables, as well as with respect to the pure jump part, in terms of an increment quotient operator related to Picard's difference operator. We prove that the well-known connection between (Z, U) and the Malliavin trace of Y still holds in the case of time delayed generators.

BSDEs without time delays and driven by Poisson random measures have already been thoroughly investigated in the literature; see [5,6] or [23]. But contrary to the case with a Gaussian basis, smoothness results in the sense of Malliavin's calculus have not been established yet in a systematic way. To the best of our knowledge, only in [7], variational differentiability of a solution of a forward–backward SDE with jumps with respect to the Brownian component is considered while differentiability with respect to the jump component is neglected.

We would like to emphasize that backward stochastic differential equations with time-delayed generators arise in financial and insurance problems dealing with pricing, hedging, risk management and optimal control; see the working paper [9]. For instance, they are encountered in the context of the optimal liquidation problem of large trader's positions. A related optimal control problem in terms of BSDEs exhibits generators in which the delayed feedback of the large trader's actions on the price dynamics take the form of a delay effect in the sense considered in this paper. As explained at the beginning of this section, Malliavin's calculus plays a fundamental role in mathematical finance and optimal control. We believe that the results concerning Malliavin's differentiability obtained in this paper are as important to describe parameter sensitivity properties of financial derivatives in this more general setting as they are in [3] for generalizing the Black–Scholes delta hedge to incomplete markets in a purely probabilistic approach via BSDEs.

This paper is structured as follows. Section 2 deals with the existence and uniqueness problem. In Section 3 we survey concepts of the canonical Lévy space and variational differentiation, and prove some technical lemmas. The main theorem concerning Malliavin smoothness of a solution, and the interpretation of the latter in terms of a Malliavin trace is proved in Section 4.

2. Existence and uniqueness of a solution

We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$, where $T < \infty$ is a finite time horizon. We assume that the filtration $\mathbb{F}$ is the natural filtration generated by a Lévy process $L := (L(t), 0 \le t \le T)$ and that $\mathcal{F}_0$ contains all sets of $\mathbb{P}$-measure zero, so that the *usual conditions* are fulfilled. As usual, by $\mathcal{B}(X)$ we denote the Borel sets of a topological space X, while λ stands for Lebesgue measure.

It is well-known that a Lévy process satisfies the Lévy–Itô decomposition

$$L(t) = at + \sigma W(t) + \int_0^t \int_{|z| \ge 1} z N(\mathrm{d}s, \mathrm{d}z) + \int_0^t \int_{0<|z|<1} z(N(\mathrm{d}s, \mathrm{d}z) - \nu(\mathrm{d}z)\mathrm{d}s),$$

for $0 \le t \le T$, with $a \in \mathbb{R}, \sigma \ge 0$. Here $W := (W(t), 0 \le t \le T)$ denotes a Brownian motion and N a random measure on $[0, T] \times (\mathbb{R} - \{0\})$, so that W and N are independent. The random measure N

$$N(t, A) = \sharp\{0 \le s \le t; \Delta L(s) \in A\}, \quad 0 \le t \le T, \ A \in \mathcal{B}(\mathbb{R} - \{0\}),$$

counts the number of jumps of a given size. It is called Poisson random measure since, for $t \in [0, T]$ and a Borel set A such that its closure does not contain zero, $N(t, A)$ is a Poisson distributed random variable. The σ-finite measure ν, defined on $\mathcal{B}(\mathbb{R} - \{0\})$, appears in the compensator $\lambda \otimes \nu$ of the random measure N. The compensated Poisson random measure (or martingale-valued measure) is denoted by $\tilde{N}(t, A) = N(t, A) - t\nu(A)$, $t \in [0, T]$, $A \in \mathcal{B}(\mathbb{R} - \{0\})$. In this paper we deal with the random measure

$$\begin{aligned}\tilde{M}(t, A) &= \int_0^t \int_A z\tilde{N}(\mathrm{d}s, \mathrm{d}z)\\ &= \int_0^t \int_A zN(\mathrm{d}s, \mathrm{d}z) - \int_0^t \int_A z\nu(\mathrm{d}z)\mathrm{d}s, \quad 0 \le t \le T,\ A \in \mathcal{B}(\mathbb{R} - \{0\}).\end{aligned}$$

It can be considered as a compensated compound Poisson random measure as, for a fixed $t \in [0, T]$ and a Borel set A the closure of which does not contain zero, $\int_0^t \int_A zN(\mathrm{d}s, \mathrm{d}z)$ is a compound Poisson distributed random variable. Finally, we introduce the σ-finite measure

$$m(A) = \int_A z^2\nu(\mathrm{d}z), \quad A \in \mathcal{B}(\mathbb{R} - \{0\}).$$

For details concerning Lévy processes, Poisson random measures and integration with respect to martingale-valued random measures we refer the reader to Chapters 2 and 4 of [4].

Let us now turn to the main subject of this paper. We study solutions $(Y, Z, U) := (Y(t), Z(t), U(t, z))_{0 \le t \le T, z \in (\mathbb{R} - \{0\})}$ of a BSDE with time delayed generator, the dynamics of which is given by

$$\begin{aligned}Y(t) = \xi &+ \int_t^T f(s, Y_s, Z_s, U_s)\mathrm{d}s\\ &- \int_t^T Z(s)\mathrm{d}W(s) - \int_t^T \int_{\mathbb{R}-\{0\}} U(s, z)\tilde{M}(\mathrm{d}s, \mathrm{d}z), \quad 0 \le t \le T.\end{aligned} \tag{2.1}$$

The generator f depends on the past values of the solution, fed back into the system with a time delay, denoted by $Y_s := (Y(s + v))_{-T \le v \le 0}$, $Z_s := (Z(s + v))_{-T \le v \le 0}$ and $U_s := (U(s + v, .))_{-T \le v \le 0}$, $0 \le s \le T$. We always set $Z(t) = U(t, .) = 0$ and $Y(t) = Y(0)$ for $t < 0$. Note that the measure $\tilde{M}$, not $\tilde{N}$, is taken to drive the jump noise. The reason for this is that we adopt the concepts of Malliavin calculus on the canonical Lévy space from [24], which is formulated in terms of multiple stochastic integrals with respect to $\tilde{M}$.

We shall work with the function spaces of the following definition.

Definition 2.1. 1. Let $L^2_{-T}(\mathbb{R})$ denote the space of measurable functions $z : [-T, 0] \to \mathbb{R}$ satisfying

$$\int_{-T}^0 |z(t)|^2\mathrm{d}t < \infty.$$

2. Let $L^2_{-T,m}(\mathbb{R})$ denote the space of product measurable functions $u : [-T, 0] \times (\mathbb{R} - \{0\}) \to \mathbb{R}$ satisfying

$$\int_{-T}^0 \int_{\mathbb{R}-\{0\}} |u(t, z)|^2 m(\mathrm{d}z)\mathrm{d}t < \infty.$$

3. Let $L^\infty_{-T}(\mathbb{R})$ denote the space of bounded, measurable functions $y : [-T, 0] \to \mathbb{R}$ such that

$$\sup_{t \in [-T,0]} |y(t)|^2 < \infty.$$

4. Let $\mathbb{L}^2(\mathbb{R})$ denote the space of $\mathcal{F}_T$-measurable random variables $\xi : \Omega \to \mathbb{R}$ which fulfill
$$\mathbb{E}\left[|\xi|^2\right] < \infty.$$
5. Let $\mathbb{H}^2_T(\mathbb{R})$ denote the space of predictable processes $Z : \Omega \times [0, T] \to \mathbb{R}$ such that
$$\mathbb{E}\left[\int_0^T |Z(t)|^2 \mathrm{d}t\right] < \infty.$$
6. Let $\mathbb{H}^2_{T,m}(\mathbb{R})$ denote the space of predictable processes $U : \Omega \times [0, T] \times (\mathbb{R} - \{0\}) \to \mathbb{R}$ satisfying
$$\mathbb{E}\left[\int_0^T \int_{\mathbb{R}-\{0\}} |U(t, z)|^2 m(\mathrm{d}z)\mathrm{d}t\right] < \infty.$$
7. Finally, let $\mathbb{S}^2_T(\mathbb{R})$ denote the space of $\mathbb{F}$-adapted, product measurable processes $Y : \Omega \times [0, T] \to \mathbb{R}$ satisfying
$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y(t)|^2\right] < \infty.$$

The spaces $\mathbb{H}^2_T(\mathbb{R})$, $\mathbb{H}^2_{T,m}(\mathbb{R})$ and $\mathbb{S}^2_T(\mathbb{R})$ are endowed with the norms
$$\|Z\|^2_{\mathbb{H}^2_T} = \mathbb{E}\left[\int_0^T \mathrm{e}^{\beta t} |Z(t)|^2 \mathrm{d}t\right],$$
$$\|U\|^2_{\mathbb{H}^2_{T,m}} = \mathbb{E}\left[\int_0^T \int_{\mathbb{R}-\{0\}} \mathrm{e}^{\beta t} |U(t, z)|^2 m(\mathrm{d}z)\mathrm{d}t\right],$$
$$\|Y\|^2_{\mathbb{S}^2_T} = \mathbb{E}\left[\sup_{t\in[0,T]} \mathrm{e}^{\beta t} |Y(t)|^2\right],$$
with some $\beta > 0$.

Predictability of Z means measurability with respect to the predictable σ-algebra, which we denote by $\mathcal{P}$, while predictability of U means measurability with respect to the product $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}-\{0\})$. In the sequel let us simply write $\mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2_m(\mathbb{R})$ instead of $\mathbb{S}^2_T(\mathbb{R}) \times \mathbb{H}^2_T(\mathbb{R}) \times \mathbb{H}^2_{T,m}(\mathbb{R})$.

We start with establishing existence and uniqueness of a solution of (2.1) under the following hypotheses:

(A1) the terminal value $\xi \in \mathbb{L}^2(\mathbb{R})$,

(A2) m is a finite measure, i.e. $\int_{\mathbb{R}-\{0\}} z^2\nu(\mathrm{d}z) < \infty$,

(A3) the generator $f : \Omega \times [0, T] \times L^\infty_{-T}(\mathbb{R}) \times L^2_{-T}(\mathbb{R}) \times L^2_{-T,m}(\mathbb{R}) \to \mathbb{R}$ is product measurable, $\mathbb{F}$-adapted and Lipschitz continuous in the sense that for a probability measure α on $([-T, 0], \mathcal{B}([-T, 0]))$ and with a constant $K > 0$
$$\begin{aligned}&|f(\omega, t, y_t, z_t, u_t) - f(\omega, t, \tilde{y}_t, \tilde{z}_t, \tilde{u}_t)|^2\\ &\quad\le K\bigg(\int_{-T}^0 |y(t+v) - \tilde{y}(t+v)|^2\alpha(\mathrm{d}v) + \int_{-T}^0 |z(t+v) - \tilde{z}(t+v)|^2\alpha(\mathrm{d}v)\\ &\qquad+ \int_{-T}^0 \int_{\mathbb{R}-\{0\}} |u(t+v, z) - \tilde{u}(t+v, z)|^2 m(\mathrm{d}z)\alpha(\mathrm{d}v)\bigg),\end{aligned}$$

holds for $\mathbb{P} \otimes \lambda$-a.e. $(\omega, t) \in \Omega \times [0, T]$, for any $(y_t, z_t, u_t), (\tilde{y}_t, \tilde{z}_t, \tilde{u}_t) \in L^{\infty}_{-T}(\mathbb{R}) \times L^2_{-T}(\mathbb{R}) \times L^2_{-T,m}(\mathbb{R})$,

(A4) $\mathbb{E}\left[\int_0^T |f(t, 0, 0, 0)|^2 \mathrm{d}t\right] < \infty$,

(A5) $f(\omega, t, ., ., .) = 0$ for $\omega \in \Omega$, $t < 0$.

For convenience, in the notation of f the dependence on ω is omitted and we write $f(t, ., ., .)$ for $f(\omega, t, ., ., .)$ etc. We remark that $f(t, 0, 0, 0)$ in (A4) should be understood as the value of the generator $f(t, y_t, z_t, u_t)$ at $y_t = z_t = u_t = 0$. We would like to point out that assumption (A5) in fact allows us to take $Y(t) = Y(0)$ and $Z(t) = U(t, .) = 0$ for $t < 0$ as a solution of (2.1). Finally, let us recall that under (A2) and for an integrand $U \in \mathbb{H}^2_m(\mathbb{R})$, the stochastic integral with respect to the martingale-valued measure $\tilde{M}$

$$\int_0^t \int_{\mathbb{R}-\{0\}} U(s, z)\tilde{M}(\mathrm{d}s, \mathrm{d}z), \quad 0 \le t \le T,$$

is well defined in the Itô sense; see Chapter 4.1 in [4].

First let us notice that for $(Y, Z, U) \in \mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2_m(\mathbb{R})$ the generator is well-defined and integrable as a consequence of

$$\begin{aligned}
\int_0^T |f(t, Y_t, Z_t, U_t)|^2 \mathrm{d}t &\le 2\int_0^T |f(t, 0, 0, 0)|^2 \mathrm{d}t + 2K\Bigg(\int_0^T \int_{-T}^0 |Y(t+v)|^2 \alpha(\mathrm{d}v)\mathrm{d}t \\
&\quad + \int_0^T \int_{-T}^0 |Z(t+v)|^2 \alpha(\mathrm{d}v)\mathrm{d}t + \int_0^T \int_{-T}^0 \int_{\mathbb{R}-\{0\}} |U(t+v, z)|^2 m(\mathrm{d}z)\alpha(\mathrm{d}v)\mathrm{d}t\Bigg) \\
&= 2\int_0^T |f(t, 0, 0, 0)|^2 \mathrm{d}t + 2K\int_{-T}^0 \int_v^{T+v} |Y(w)|^2 \mathrm{d}w\alpha(\mathrm{d}v) \\
&\quad + 2K\int_{-T}^0 \int_v^{T+v} |Z(w)|^2 \mathrm{d}w\alpha(\mathrm{d}v) \\
&\quad + 2K\int_{-T}^0 \int_v^{T+v} \int_{\mathbb{R}-\{0\}} |U(w, z)|^2 m(\mathrm{d}z)\mathrm{d}w\alpha(\mathrm{d}v) \\
&\le 2\int_0^T |f(t, 0, 0, 0)|^2 \mathrm{d}t + 2K\Bigg(T \sup_{w\in[0,T]} |Y(w)|^2 \\
&\quad + \int_0^T |Z(w)|^2 \mathrm{d}w + \int_0^T \int_{\mathbb{R}-\{0\}} |U(w, z)|^2 m(\mathrm{d}z)\mathrm{d}w\Bigg) < \infty, \quad \mathbb{P}\text{-a.s.,}
\end{aligned} \tag{2.2}$$

where we apply (A3), Fubini's theorem, use the assumption that $Z(t) = U(t, .) = 0$ and $Y(t) = Y(0)$ for $t < 0$ and the fact that the measure α is a probability measure.

The main theorem of this section is an extension of Theorem 2.1 from [10]. Although the extension is quite natural, the proof is given for completeness and convenience of the reader. The key result follows from the following a priori estimates.

Lemma 2.1. *Let* $(Y, Z, U), (\tilde{Y}, \tilde{Z}, \tilde{U}) \in \mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2_m(\mathbb{R})$ *denote solutions of* (2.1) *with corresponding parameters* (ξ, f) *and* $(\tilde{\xi}, \tilde{f})$ *which satisfy the assumptions* (A1)–(A5). *Then the* *following inequalities hold*

$$\|Z-\tilde{Z}\|^2_{\mathbb{H}^2}+\|U-\tilde{U}\|^2_{\mathbb{H}^2_m}\le \mathrm{e}^{\beta T}\mathbb{E}\left[\left|\xi-\tilde{\xi}\right|^2\right]$$
$$+\frac{1}{\beta}\mathbb{E}\left[\int_0^T \mathrm{e}^{\beta t}|f(t,Y_t,Z_t,U_t)-\tilde{f}(t,\tilde{Y}_t,\tilde{Z}_t,\tilde{U}_t)|^2\mathrm{d}t\right], \tag{2.3}$$
$$\|Y-\tilde{Y}\|^2_{\mathbb{S}^2}\le 8\mathrm{e}^{\beta T}\mathbb{E}\left[\left|\xi-\tilde{\xi}\right|^2\right]$$
$$+8T\mathbb{E}\left[\int_0^T \mathrm{e}^{\beta t}|f(t,Y_t,Z_t,U_t)-\tilde{f}(t,\tilde{Y}_t,\tilde{Z}_t,\tilde{U}_t)|^2\mathrm{d}t\right]. \tag{2.4}$$

Proof. The inequality (2.3) follows by a straightforward extension of Lemma 3.2.1 from [14], by only adding an additional stochastic integral with respect to $\tilde{M}$. In order to prove the second inequality, first notice that for $t\in[0,T]$

$$Y(t)-\tilde{Y}(t)=\mathbb{E}\left[\xi-\tilde{\xi}+\int_t^T (f(s,Y_s,Z_s,U_s)-\tilde{f}(s,\tilde{Y}_s,\tilde{Z}_s,\tilde{U}_s))\mathrm{d}s|\mathcal{F}_t\right],$$

and

$$\mathrm{e}^{\frac{\beta}{2}t}|Y(t)-\tilde{Y}(t)|\le \mathrm{e}^{\frac{\beta}{2}T}\mathbb{E}\left[|\xi-\tilde{\xi}||\mathcal{F}_t\right]$$
$$+\mathbb{E}\left[\int_0^T \mathrm{e}^{\frac{\beta}{2}s}|f(s,Y_s,Z_s,U_s)-\tilde{f}(s,\tilde{Y}_s,\tilde{Z}_s,\tilde{U}_s)|\mathrm{d}s|\mathcal{F}_t\right],$$

hold $\mathbb{P}$-a.s. Doob's martingale inequality and Cauchy–Schwarz' inequality yield the second estimate. The reader may also consult Proposition 2.2 in [5] or Proposition 3.3 in [6], where similar estimates for BSDEs with jumps are derived. □

Theorem 2.1. *Assume that* (A1)–(A5) *hold. For a sufficiently small time horizon T or for a sufficiently small Lipschitz constant K of the generator f, more precisely if for some $\beta>0$*

$$\delta(T,K,\beta,\alpha):=\left(8T+\frac{1}{\beta}\right)K\int_{-T}^0 \mathrm{e}^{-\beta v}\alpha(\mathrm{d}v)\max\{1,T\}<1,$$

the backward stochastic differential equation (2.1) has a unique solution $(Y,Z,U)\in\mathbb{S}^2(\mathbb{R})\times\mathbb{H}^2(\mathbb{R})\times\mathbb{H}^2_m(\mathbb{R})$.

Proof. We follow the classical Picard type iteration scheme (see Theorem 2.1 in [12] or Theorem 3.2.1 in [14]) to prove existence and uniqueness of a solution.

Let $Y^0(t)=Z^0(t)=U^0(t,z)=0$, $(t,z)\in[0,T]\times(\mathbb{R}-\{0\})$.

Step (1) We show that the recursive definition

$$Y^{n+1}(t)=\xi+\int_t^T f(s,Y^n_s,Z^n_s,U^n_s)\mathrm{d}s-\int_t^T Z^{n+1}(s)\mathrm{d}W(s)$$
$$-\int_t^T\int_{\mathbb{R}-\{0\}} U^{n+1}(s,z)\tilde{M}(\mathrm{d}s,\mathrm{d}z),\quad 0\le t\le T, \tag{2.5}$$

makes sense. More precisely, we show that given $(Y^n,Z^n,U^n)\in\mathbb{S}^2(\mathbb{R})\times\mathbb{H}^2(\mathbb{R})\times\mathbb{H}^2_m(\mathbb{R})$, Eq. (2.5) has a unique solution $(Y^{n+1},Z^{n+1},U^{n+1})\in\mathbb{S}^2(\mathbb{R})\times\mathbb{H}^2(\mathbb{R})\times\mathbb{H}^2_m(\mathbb{R})$.

Applying inequality (2.2), we can conclude that

$$\mathbb{E}\left[\int_0^T |f(t, Y_t^n, Z_t^n, U_t^n)|^2 \mathrm{d}t\right] \leq 2\mathbb{E}\left[\int_0^T |f(t, 0, 0, 0)|^2 \mathrm{d}t\right] + 2K\left(T\|Y^n\|_{\mathbb{S}^2} + \|Z^n\|_{\mathbb{H}^2} + \|U^n\|_{\mathbb{H}^2_m}\right) < \infty.$$

As in the case of BSDEs without time delays, the martingale representation, see Theorem 13.49 in [13], provides a unique process $Z^{n+1} \in \mathbb{H}^2(\mathbb{R})$ and a unique predictable process $\bar{U}^{n+1}$ satisfying

$$\mathbb{E}\left[\int_0^T \int_{\mathbb{R}-\{0\}} |\bar{U}^{n+1}(t, z)|^2 \nu(\mathrm{d}z)\mathrm{d}t\right] < \infty,$$

so that

$$\xi + \int_0^T f(t, Y_t^n, Z_t^n, U_t^n)\mathrm{d}t = \mathbb{E}\left[\xi + \int_0^T f(t, Y_t^n, Z_t^n, U_t^n)\mathrm{d}t\right] + \int_0^T Z^{n+1}(t)\mathrm{d}W(t) + \int_0^T \int_{\mathbb{R}-\{0\}} \bar{U}^{n+1}(t, z)\tilde{N}(\mathrm{d}t, \mathrm{d}z), \quad \mathbb{P}\text{-a.s.}$$

For $(t, z) \in [0, T] \times (\mathbb{R} - \{0\})$ we get $U^{n+1}(t, z) = \frac{\bar{U}^{n+1}(t,z)}{z} \in \mathbb{H}^2_m(\mathbb{R})$, and have the required representation

$$\xi + \int_0^T f(t, Y_t^n, Z_t^n, U_t^n)\mathrm{d}t = \mathbb{E}\left[\xi + \int_0^T f(t, Y_t^n, Z_t^n, U_t^n)\mathrm{d}t\right] + \int_0^T Z^{n+1}(t)\mathrm{d}W(t) + \int_0^T \int_{\mathbb{R}-\{0\}} U^{n+1}(t, z)(t)\tilde{M}(\mathrm{d}t, \mathrm{d}z), \quad \mathbb{P}\text{-a.s.}$$

Finally, we take Y^{n+1} as a progressively measurable, càdlàg modification of

$$Y^{n+1}(t)(\omega) = \mathbb{E}\left[\xi + \int_t^T f(s, Y_s^n, Z_s^n, U_s^n)\mathrm{d}s|\mathcal{F}_t\right], \quad \omega \in \Omega,\ t \in [0, T].$$

Similarly as in Lemma 2.1, Doob's martingale inequality, Cauchy–Schwarz' inequality and the estimates (2.2) yield that $Y^{n+1} \in \mathbb{S}^2(\mathbb{R})$.

Step (2) We prove the convergence of the sequence (Y^n, Z^n, U^n) in $\mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2_m(\mathbb{R})$. The estimates (2.3) and (2.4) provide the inequality

$$\begin{aligned} &\|Y^{n+1} - Y^n\|^2_{\mathbb{S}^2} + \|Z^{n+1} - Z^n\|^2_{\mathbb{H}^2} + \|U^{n+1} - U^n\|^2_{\mathbb{H}^2_m} \\ &\leq \left(8T + \frac{1}{\beta}\right)\mathbb{E}\left[\int_0^T \mathrm{e}^{\beta t}|f(t, Y_t^n, Z_t^n, U_t^n) - f(t, Y_t^{n-1}, Z_t^{n-1}, U_t^{n-1})|^2 \mathrm{d}t\right]. \end{aligned} \tag{2.6}$$

By applying the Lipschitz condition (A3), Fubini's theorem, changing variables and using the assumption $\forall n \geq 0 Y^n(s) = Y^n(0)$ and $Z^n(s) = U^n(s, .) = 0$ for $s < 0$, we can derive

$$\begin{aligned} &\mathbb{E}\left[\int_0^T \mathrm{e}^{\beta t}|f(t, Y_t^n, Z_t^n, U_t^n) - f(t, Y_t^{n-1}, Z_t^{n-1}, U_t^{n-1})|^2 \mathrm{d}t\right] \\ &\leq K\mathbb{E}\Bigg[\int_0^T \mathrm{e}^{\beta t}\int_{-T}^0 |Y^n(t+v) - Y^{n-1}(t+v)|^2 \alpha(\mathrm{d}v)\mathrm{d}t \end{aligned}$$

$$
\begin{aligned}
&\quad + \int_0^T \mathrm{e}^{\beta t} \int_{-T}^0 |Z^n(t+v) - Z^{n-1}(t+v)|^2 \alpha(\mathrm{d}v)\mathrm{d}t \\
&\quad + \int_0^T \mathrm{e}^{\beta t} \int_{-T}^0 \int_{\mathbb{R}-\{0\}} |U^n(t+v,z) - U^{n-1}(t+v,z)|^2 m(\mathrm{d}z)\alpha(\mathrm{d}v)\mathrm{d}t \Bigg] \\
&= K\mathbb{E}\Bigg[\int_{-T}^0 \mathrm{e}^{-\beta v} \int_0^T \mathrm{e}^{\beta(t+v)} |Y^n(t+v) - Y^{n-1}(t+v)|^2 \mathrm{d}t\alpha(\mathrm{d}v) \\
&\quad + \int_{-T}^0 \mathrm{e}^{-\beta v} \int_0^T \mathrm{e}^{\beta(t+v)} |Z^n(t+v) - Z^{n-1}(t+v)|^2 \mathrm{d}t\alpha(\mathrm{d}v) \\
&\quad + \int_{-T}^0 \mathrm{e}^{-\beta v} \int_0^T \int_{\mathbb{R}-\{0\}} \mathrm{e}^{\beta(t+v)} |U^n(t+v,z) - U^{n-1}(t+v,z)|^2 m(\mathrm{d}z)\mathrm{d}t\alpha(\mathrm{d}v) \Bigg] \\
&= K\mathbb{E}\Bigg[\int_{-T}^0 \mathrm{e}^{-\beta v} \int_v^{T+v} \mathrm{e}^{\beta w} |Y^n(w) - Y^{n-1}(w)|^2 \mathrm{d}w\alpha(\mathrm{d}v) \\
&\quad + \int_{-T}^0 \mathrm{e}^{-\beta v} \int_v^{T+v} \mathrm{e}^{\beta w} |Z^n(w) - Z^{n-1}(w)|^2 \mathrm{d}w\alpha(\mathrm{d}v) \\
&\quad + \int_{-T}^0 \mathrm{e}^{-\beta v} \int_v^{T+v} \int_{\mathbb{R}-\{0\}} \mathrm{e}^{\beta w} |U^n(w,z) - U^{n-1}(w,z)|^2 m(\mathrm{d}z)\mathrm{d}w\alpha(\mathrm{d}v) \Bigg] \\
&\le K \int_{-T}^0 \mathrm{e}^{-\beta v}\alpha(\mathrm{d}v) \left(T\|Y^n - Y^{n-1}\|^2_{\mathbb{S}^2} + \|Z^n - Z^{n-1}\|^2_{\mathbb{H}^2} + \|U^n - U^{n-1}\|^2_{\mathbb{H}^2_m} \right).
\end{aligned}
\tag{2.7}
$$

From (2.6) and (2.7), we obtain

$$
\begin{aligned}
&\|Y^{n+1} - Y^n\|^2_{\mathbb{S}^2} + \|Z^{n+1} - Z^n\|^2_{\mathbb{H}^2} + \|U^{n+1} - U^n\|^2_{\mathbb{H}^2_m} \\
&\quad \le \delta(T,K,\beta,\alpha) \left(\|Y^n - Y^{n-1}\|^2_{\mathbb{S}^2} + \|Z^n - Z^{n-1}\|^2_{\mathbb{H}^2} + \|U^n - U^{n-1}\|^2_{\mathbb{H}^2_m} \right),
\end{aligned}
\tag{2.8}
$$

with

$$
\delta(T,K,\beta,\alpha) = \left(8T + \frac{1}{\beta} \right) K \int_{-T}^0 \mathrm{e}^{-\beta v}\alpha(\mathrm{d}v) \max\{1,T\}.
$$

For $\beta = \frac{1}{T}$ we have

$$
\delta(T,K,\beta,\alpha) \le 9TK\mathrm{e}\max\{1,T\}.
$$

For sufficiently small T or sufficiently small K, the inequality (2.8) provides a unique limit $(Y,Z,U) \in \mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2_m(\mathbb{R})$ of the converging sequence $(Y^n,Z^n,U^n)_{n\in\mathbb{N}}$, which satisfies the fixed point equation

$$
Y(t) = \mathbb{E}\left[\xi + \int_t^T f(s,Y_s,Z_s,U_s)\mathrm{d}s | \mathcal{F}_t \right], \quad \mathbb{P}\text{-a.s., } 0 \le t \le T.
$$

Step (3) We define the solution component $\bar{Y}$ of (4.1) as a progressively measurable, càdlàg modification of

$$\bar{Y}(t)(\omega) = \mathbb{E}\left[\xi + \int_t^T f(s, Y_s, Z_s, U_s)\mathrm{d}s|\mathcal{F}_t\right], \quad \omega \in \Omega,\ t \in [0, T],$$

where (Y, Z, U) is the limit constructed in Step (2). □

We point out that in general under the assumptions (A1)–(A5), existence and uniqueness may fail to hold for bigger time horizon T or bigger Lipschitz constant K. See [10] for examples. However, for some special classes of generators existence and uniqueness may be proved for an arbitrary time horizon and for arbitrary global Lipschitz constants. These include generators independent of y with a delay measure α supported on $[-\gamma, 0]$ with a sufficiently small time delay γ, following Theorem 2.2 in [10], or generators considered in [8] consisting of separate components in z and u, following Theorem 1 in [8].

3. Malliavin's calculus for canonical Lévy processes

There are various ways to develop Malliavin's calculus for Lévy processes. In this paper we adopt the approach from [24] based on a chaos decomposition in terms of multiple stochastic integrals with respect to the random measure $\tilde{M}$. In this setting, we will construct a suitable canonical space, on which a variational derivative with respect to the pure jump part of a Lévy process can be computed in a pathwise sense.

In this section we give an overview of the approach of Malliavin's calculus on canonical Lévy space according to [24] (see [24] for details). We then prove some technical results concerning the commutation of integration and variational differentiation, which are needed in the next section.

We assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is the product of two canonical spaces $(\Omega_W \times \Omega_N, \mathcal{F}_W \times \mathcal{F}_N, \mathbb{P}_W \times \mathbb{P}_N)$, and the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ the canonical filtration completed for $\mathbb{P}$. The space $(\Omega_W, \mathcal{F}_W, \mathbb{P}_W)$ is the usual canonical space for a one-dimensional Brownian motion, with the space of continuous functions on $[0, T]$, the σ-algebra generated by the topology of uniform convergence and Wiener measure. The canonical representation for a pure jump Lévy process $(\Omega_N, \mathcal{F}_N, \mathbb{P}_N)$ we use is based on a fixed partition $(S_k)_{k\geq 1}$ of $\mathbb{R} - \{0\}$ such that $0 < \nu(S_k) < \infty, k \geq 1$. Accordingly, it is given by the product space $\bigotimes_{k\geq 1}(\Omega_N^k, \mathcal{F}_N^k, \mathbb{P}_N^k)$ of spaces of compound Poisson processes on $[0, T]$ with intensities $\nu(S_k)$ and jump size distributions supported on $S_k, k \geq 1$. Since trajectories of compound Poisson processes can be described by finite families $((t_1, z_1), \ldots, (t_n, z_n))$, where $(t_1, \ldots, t_n)$ denotes the jump times and $(z_1, \ldots, z_n)$ the corresponding sizes of jumps, one can take $\Omega_N^k = \bigcup_{n\geq 0}([0, T] \times (\mathbb{R} - \{0\}))^n$, with $([0, T] \times (\mathbb{R} - \{0\}))^0$ representing an empty sequence, the σ-algebra $\mathcal{F}_N^k = \bigvee_{n\geq 0} \mathcal{B}(([0, T] \times (\mathbb{R} - \{0\}))^n)$, and the measure $\mathbb{P}_N^k$ defined in such a way that for $B = \cup_{n\geq 0} B_n$, $B_n \in \mathcal{B}(([0, T] \times (\mathbb{R} - \{0\}))^n)$, we have

$$\mathbb{P}_N^k(B) = \mathrm{e}^{-\nu(S_k)T} \sum_{n=0}^{\infty} \frac{(\nu(S_k))^n \left(\mathrm{d}t \otimes \frac{\nu \mathbf{1}\{S_k\}}{\nu(S_k)}\right)^{\otimes n}(B_n)}{n!}.$$

Now consider the finite measure q defined on $[0, T] \times \mathbb{R}$ by

$$q(E) = \int_{E(0)} \mathrm{d}t + \int_{E'} z^2 \nu(\mathrm{d}z)\mathrm{d}t, \quad E \in \mathcal{B}([0, T] \times \mathbb{R}),$$

where $E(0) = \{t \in [0, T]; (t, 0) \in E\}$ and $E' = E - E(0)$, and the random measure Q on $[0, T] \times \mathbb{R}$

$$Q(E) = \int_{E(0)} \mathrm{d}W(t) + \int_{E'} z\tilde{N}(\mathrm{d}t, \mathrm{d}z), \quad E \in \mathcal{B}([0, T] \times \mathbb{R}).$$

For $n \in \mathbb{N}$ and a simple function $h_n = \mathbf{1}_{E_1 \times \cdots \times E_n}$, with pairwise disjoints sets $E_1, \ldots, E_n \in \mathcal{B}([0, T] \times \mathbb{R})$, a multiple two-parameter integral with respect to the random measure Q

$$I_n(h_n) = \int_{([0,T]\times\mathbb{R})^n} h((t_1, z_1), \ldots, (t_n, z_n))Q(\mathrm{d}t_1, \mathrm{d}z_1) \cdots Q(\mathrm{d}t_n, \mathrm{d}z_n)$$

can be defined as

$$I_n(h_n) = Q(E_1) \ldots Q(E_n).$$

The integral can be extended to the space $L^2_{T,q,n}(\mathbb{R})$ of product measurable, deterministic functions $h : ([0, T] \times \mathbb{R})^n \to \mathbb{R}$ satisfying

$$\|h\|^2_{L^2_{T,q,n}} = \int_{([0,T]\times\mathbb{R})^n} |h_n((t_1, z_1), \ldots, (t_n, z_n))|^2 q(\mathrm{d}t_1, \mathrm{d}z_1) \cdots q(\mathrm{d}t_n, \mathrm{d}z_n) < \infty.$$

The chaotic decomposition property yields that any $\mathcal{F}$-measurable square integrable random variable H on the canonical space has a unique representation

$$H = \sum_{n=0}^{\infty} I_n(h_n), \quad \mathbb{P}\text{-a.s.}, \tag{3.1}$$

with functions $h_n \in L^2_{T,q,n}(\mathbb{R})$ that are symmetric in the n pairs (t_i, z_i), $1 \le i \le n$. Moreover,

$$\mathbb{E}\left[H^2\right] = \sum_{n=0}^{\infty} n!\|h_n\|^2_{L^2_{T,q,n}}. \tag{3.2}$$

In this setting it is possible to study two-parameter annihilation operators (Malliavin derivatives) and creation operators (Skorokhod integrals).

Definition 3.1. 1. Let $\mathbb{D}^{1,2}(\mathbb{R})$ denote the space of $\mathbb{F}$-measurable random variables $H \in \mathbb{L}^2(\mathbb{R})$ with the representation $H = \sum_{n=0}^{\infty} I_n(h_n)$ satisfying

$$\sum_{n=1}^{\infty} nn!\|h_n\|^2_{L^2_{T,q,n}} < \infty.$$

2. The Malliavin derivative $DH : \Omega \times [0, T] \times \mathbb{R} \to \mathbb{R}$ of a random variable $H \in \mathbb{D}^{1,2}(\mathbb{R})$ is a stochastic process defined by

$$D_{t,z}H = \sum_{n=1}^{\infty} nI_{n-1}(h_n((t, z), \cdot)), \quad \text{valid for } q\text{-a.e. } (t, z) \in [0, T] \times \mathbb{R},\ \mathbb{P}\text{-a.s.}$$

3. Let $\mathbb{L}^{1,2}(\mathbb{R})$ denote the space of product measurable and $\mathbb{F}$-adapted processes $G : \Omega \times [0, T] \times \mathbb{R} \to \mathbb{R}$ satisfying

$$\mathbb{E}\left[\int_{[0,T]\times\mathbb{R}} |G(s, y)|^2 q(\mathrm{d}s, \mathrm{d}y)\right] < \infty,$$

$$G(s, y) \in \mathbb{D}^{1,2}(\mathbb{R}), \quad \text{for } q\text{-a.e. } (s, y) \in [0, T] \times \mathbb{R},$$

$$\mathbb{E}\left[\int_{([0,T]\times\mathbb{R})^2} |D_{t,z}G(s, y)|^2 q(\mathrm{d}s, \mathrm{d}y)q(\mathrm{d}t, \mathrm{d}z)\right] < \infty.$$

In terms of the components of the representation of $G(s,y)=\sum_{n=0}^{\infty} I_n((g_n(s,y),.))$, for q-a.e. $(s,y)\in[0,T]\times\mathbb{R}$, the above conditions are equivalent to

$$\sum_{n=1}^{\infty}(n+1)(n+1)!\|\hat{g}_n\|^2_{L^2_{T,q,n+1}}<\infty,$$

where $\hat{g}_n$ denotes the symmetrization of g_n with respect to all $n+1$ pairs of variables. The space $\mathbb{L}^{1,2}(\mathbb{R})$ is a Hilbert space endowed with the norm

$$\|G\|^2_{\mathbb{L}^{1,2}}=\mathbb{E}\left[\int_{[0,T]\times\mathbb{R}}|G(s,y)|^2q(\mathrm{d}s,\mathrm{d}y)\right] +\mathbb{E}\left[\int_{([0,T]\times\mathbb{R})^2}|D_{t,z}G(s,y)|^2q(\mathrm{d}s,\mathrm{d}y)q(\mathrm{d}t,\mathrm{d}z)\right].$$

4. The Skorokhod integral with respect to the random measure Q of a process $G:\Omega\times[0,T]\times\mathbb{R}\to\mathbb{R}$ with the representation $G(s,y)=\sum_{n=0}^{\infty} I_n((g_n(s,y),.))$, for q-a.e. $(s,y)\in[0,T]\times\mathbb{R}$, satisfying

$$\sum_{n=0}^{\infty}(n+1)!\|\hat{g}_n\|^2_{L^2_{T,q,n+1}}<\infty,$$

is defined as

$$\int_{[0,T]\times\mathbb{R}}G(s,y)Q(\mathrm{d}s,\mathrm{d}y)=\sum_{n=0}^{\infty}I_{n+1}(\hat{g}_n),\quad \mathbb{P}\text{-a.s.}$$

The following practical rules of differentiation hold. Consider a random variable H defined on $\Omega_W\times\Omega_N$. The derivative $D_{t,0}H$ is with respect to the Brownian motion component of the Lévy process, and we can apply classical Malliavin's calculus for Hilbert space-valued random variables. If for $\mathbb{P}^N$-a.e. $\omega_N\in\Omega_N$ the random variable $H(.,\omega_N)$ is differentiable in the sense of classical Malliavin's calculus, then we have the relation

$$D_{t,0}H(\omega_W,\omega_N)=D_tH(.,\omega_N)(\omega_W),\quad \lambda\text{-a.e. } t\in[0,T],\ \mathbb{P}^W\times\mathbb{P}^N\text{-a.s.},\tag{3.3}$$

where D_t denotes the classical Malliavin derivative on the canonical Brownian space; see Proposition 3.5 in [24]. In order to define $D_{t,z}F$ for $z\neq 0$, which is a derivative with respect to the pure jump part of the Lévy process, consider the following increment quotient operator

$$\Psi_{t,z}H(\omega_W,\omega_N)=\frac{H(\omega_W,\omega_N^{t,z})-H(\omega_W,\omega_N)}{z},\tag{3.4}$$

where $\omega_N^{t,z}$ transforms a family $\omega_N=((t_1,z_1),(t_2,z_2),\ldots)\in\Omega_N$ into a new family $\omega_N^{t,z}=((t,z),(t_1,z_1),(t_2,z_2),\ldots)\in\Omega_N$, by adding a jump of size z at time t into the trajectory. According to Propositions 5.4 and 5.5 in [24], for $H\in\mathbb{L}^2(\mathbb{R})$ such that $\mathbb{E}\left[\int_0^T\int_{\mathbb{R}-\{0\}}\left|\Psi_{t,z}H\right|^2 m(\mathrm{d}z)\mathrm{d}t\right]<\infty$ we have the relation

$$D_{t,z}H=\Psi_{t,z}H,\quad \text{for } \lambda\otimes m\text{-a.e. } (t,z)\in[0,T]\times(\mathbb{R}-\{0\}),\ \mathbb{P}\text{-a.s.}\tag{3.5}$$

The operator (3.4) is closely related to Picard's difference operator, introduced in [22], which is just the numerator of (3.4). It is possible to define Malliavin's derivative for pure jump processes in such a way that it coincides with Picard's difference operator; see [11]. We point out once again that we adopt the approach of [24], and define multiple two-parameter integrals with respect to the random measure $\tilde{M}$ and not with respect to $\tilde{N}$, to obtain differentiation rules (3.3) and (3.5).

We now discuss some technical problems arising in the next section in the context of the main theorem of this paper. The subsequent lemmas are extensions of classical Malliavin differentiation rules to the setting of the canonical Lévy space.

Lemma 3.1. *Assume that $H \in \mathbb{D}^{1,2}(\mathbb{R})$. Then, for $0 \leq s \leq T$, $\mathbb{E}[H|\mathcal{F}_s] \in \mathbb{D}^{1,2}(\mathbb{R})$ and*

$$D_{t,z}\mathbb{E}[H|\mathcal{F}_s] = \mathbb{E}\left[D_{t,z}H|\mathcal{F}_s\right]\mathbf{1}\{t \leq s\}, \quad \text{for } q\text{-a.e. } (t,z) \in [0,T] \times \mathbb{R}, \mathbb{P}\text{-a.s.}$$

Proof. The proof is a straightforward extension of the proof of Proposition 1.2.8 from [19]. Details are left to the reader. □

We next provide a proof of the commutation of Lebesgue's integration and variational differentiability, which is commonly used.

Lemma 3.2. *Let $G : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$ be a product measurable and $\mathbb{F}$-adapted process, η on $[0,T] \times \mathbb{R}$ a finite measure, so that the conditions*

$$\begin{aligned}
&\mathbb{E}\left[\int_{[0,T]\times\mathbb{R}} |G(s,y)|^2 \eta(\mathrm{d}s,\mathrm{d}y)\right] < \infty,\\
&G(s,y) \in \mathbb{D}^{1,2}(\mathbb{R}), \quad \text{for } \eta\text{-a.e. } (s,y) \in [0,T] \times \mathbb{R},\\
&\mathbb{E}\left[\int_{([0,T]\times\mathbb{R})^2} |D_{t,z}G(s,y)|^2 \eta(\mathrm{d}s,\mathrm{d}y) q(\mathrm{d}t,\mathrm{d}z)\right] < \infty
\end{aligned} \tag{3.6}$$

are satisfied. Then $\int_{[0,T]\times\mathbb{R}} G(s,y)\eta(\mathrm{d}s,\mathrm{d}y) \in \mathbb{D}^{1,2}(\mathbb{R})$ and the differentiation rule

$$D_{t,z}\int_{[0,T]\times\mathbb{R}} G(s,y)\eta(\mathrm{d}s,\mathrm{d}y) = \int_{[0,T]\times\mathbb{R}} D_{t,z}G(s,y)\eta(\mathrm{d}s,\mathrm{d}y)$$

holds for q-a.e. $(t,z) \in [0,T] \times \mathbb{R}$, $\mathbb{P}$-a.s.

Proof. As for η-a.e. $(s,y) \in [0,T] \times \mathbb{R}$ the random variable $G(s,y)$ is $\mathcal{F}_s$-measurable and square integrable, the chaotic decomposition property on the canonical space (3.1) provides a unique representation

$$G(s,y) = \sum_{n=0}^{\infty} I_n(g_n((s,y),.)), \quad \eta\text{-a.e. } (s,y) \in [0,T] \times \mathbb{R},\ \mathbb{P}\text{-a.s.}$$

By part 3 of Definition 3.1, the assumptions (3.6) yield

$$\int_{[0,T]\times\mathbb{R}} \sum_{n=1}^{\infty} nn! \|g_n((s,y),.)\|^2_{L^2_{T,q,n}} \eta(\mathrm{d}s,\mathrm{d}y) < \infty. \tag{3.7}$$

For $N \in \mathbb{N}$ let G^N be a measurable version of the partial sum of the first $N+1$ components given by

$$G^N(s,y) = \sum_{n=0}^{N} I_n(g_n((s,y),.)), \quad \eta\text{-a.e. } (s,y) \in [0,T] \times \mathbb{R},\ \mathbb{P}\text{-a.s.}$$

We first prove that $\int_{[0,T]\times\mathbb{R}} G^N(s,y)\eta(\mathrm{d}s,\mathrm{d}y) \in \mathbb{D}^{1,2}(\mathbb{R})$ and the claimed differentiation rule holds.

By canonical extension arguments, a Fubini type property holds for any single chaos component, and therefore

$$\begin{aligned}\int_{[0,T]\times\mathbb{R}} G^N(s,y)\eta(\mathrm{d}s,\mathrm{d}y) &= \int_{[0,T]\times\mathbb{R}} \sum_{n=0}^{N} \int_{([0,T]\times\mathbb{R})^n} g_n((s,y),(t_1,z_1),\ldots,(t_n,z_n)) \\ &\quad \times Q(\mathrm{d}t_1,\mathrm{d}z_1)\ldots Q(\mathrm{d}t_n,\mathrm{d}z_n)\eta(\mathrm{d}s,\mathrm{d}y) \\ &= \sum_{n=0}^{N} \int_{([0,T]\times\mathbb{R})^n} \int_{[0,T]\times\mathbb{R}} g_n((s,y),(t_1,z_1),\ldots,(t_n,z_n)) \\ &\quad \times \eta(\mathrm{d}s,\mathrm{d}y) Q(\mathrm{d}t_1,\mathrm{d}z_1)\ldots Q(\mathrm{d}t_n,\mathrm{d}z_n) \\ &= \sum_{n=0}^{N} I_n(h_n) := H^N, \end{aligned} \tag{3.8}$$

with

$$h_n((t_1,z_1),\ldots,(t_n,z_n)) = \int_{[0,T]\times\mathbb{R}} g_n((s,y),(t_1,z_1),\ldots,(t_n,z_n))\eta(\mathrm{d}s,\mathrm{d}y)$$

for $(t_1,z_1),\ldots,(t_n,z_n) \in ([0,T]\times\mathbb{R})^n$. Notice that by Cauchy–Schwarz' inequality, finiteness of η, the assumption (3.7) and Fubini's theorem we obtain

$$\sum_{n=1}^{\infty} n n! \|h_n\|^2_{L^2_{T,q,n}} < \infty. \tag{3.9}$$

For any $N \in \mathbb{N}$, we have that $H^N \in \mathbb{D}^{1,2}(\mathbb{R})$, hence $\int_{[0,T]\times\mathbb{R}} G^N(s,y)\eta(\mathrm{d}s,\mathrm{d}y) \in \mathbb{D}^{1,2}(\mathbb{R})$, and, by linearity and definition

$$\begin{aligned} D_{t,z}H^N &= D_{t,z}\int_{[0,T]\times\mathbb{R}} G^N(s,y)\eta(\mathrm{d}s,\mathrm{d}y) \\ &= \sum_{n=1}^{N} n \int_{[0,T]\times\mathbb{R}} g_n((s,y),(t,z),(t_2,z_2,),\ldots,(t_n,z_n))\eta(\mathrm{d}s,\mathrm{d}y) \\ &\quad \times Q(\mathrm{d}t_2,\mathrm{d}z_2)\ldots Q(\mathrm{d}t_n,\mathrm{d}z_n) \\ &= \int_{[0,T]\times\mathbb{R}} \sum_{n=1}^{N} n g_n((s,y),(t,z),(t_2,z_2,),\ldots,(t_n,z_n)) \\ &\quad \times Q(\mathrm{d}t_2,\mathrm{d}z_2)\ldots Q(\mathrm{d}t_n,\mathrm{d}z_n)\eta(\mathrm{d}s,\mathrm{d}y) \\ &= \int_{[0,T]\times\mathbb{R}} D_{t,z}G^N(s,y)\eta(\mathrm{d}s,\mathrm{d}y), \quad q\text{-a.e. } (t,z)\in[0,T]\times\mathbb{R}. \end{aligned}$$

The differentiation rule is proved for G^N.
Finally, by (3.9) we have

$$\begin{aligned} &\mathbb{E}\left[|H^N - H^M|^2\right] + \int_{[0,T]\times\mathbb{R}} \mathbb{E}\left[|D_{t,z}H^N - D_{t,z}H^M|^2\right] q(\mathrm{d}t,\mathrm{d}z) \\ &\quad \le \sum_{n=N+1}^{M} n n! \|h_n\|^2_{L^2_{T,q,n}} \to 0, \quad N,M\to\infty. \end{aligned}$$

By closeability of the operator D we conclude that the unique limit H is Malliavin differentiable. The convergences $G^N \to G\,\mathbb{P}\otimes\eta$-a.e. and $DG^N \to DG\,\mathbb{P}\otimes\eta\otimes q$-a.e. together with Lebesgue's dominated convergence theorem, justified by the first and third assumption in (3.6), give

$$\mathbb{E}\left[\left|\int_{[0,T]\times\mathbb{R}} G^N(s,y)\eta(\mathrm{d}s,\mathrm{d}y) - \int_{[0,T]\times\mathbb{R}} G(s,y)\eta(\mathrm{d}s,\mathrm{d}y)\right|^2\right] + \int_{[0,T]\times\mathbb{R}} \mathbb{E}\left[\left|\int_{[0,T]\times\mathbb{R}} D_{t,z}G^N(s,y)\eta(\mathrm{d}s,\mathrm{d}y) - \int_{[0,T]\times\mathbb{R}} D_{t,z}G(s,y)\eta(\mathrm{d}s,\mathrm{d}y)\right|^2\right] \times q(\mathrm{d}t,\mathrm{d}z) \to 0.$$

This implies the claimed equation. □

We finally discuss the commutation relation of the Skorokhod stochastic integral with the variational derivative.

Lemma 3.3. *Assume that $G : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$ is a predictable process and $\mathbb{E}\left[\int_{[0,T]\times\mathbb{R}} |G(s,y)|^2 q(\mathrm{d}s,\mathrm{d}y)\right] < \infty$ holds. Then*

$$G \in \mathbb{L}^{1,2}(\mathbb{R}) \quad \textit{if and only if} \quad \int_{[0,T]\times\mathbb{R}} G(s,y)Q(\mathrm{d}s,\mathrm{d}y) \in \mathbb{D}^{1,2}(\mathbb{R}).$$

Moreover, if $\int_{[0,T]\times\mathbb{R}} G(s,y)Q(\mathrm{d}s,\mathrm{d}y) \in \mathbb{D}^{1,2}(\mathbb{R})$ then, for q-a.e. $(t,z) \in [0,T]\times\mathbb{R}$,

$$D_{t,z}\int_{[0,T]\times\mathbb{R}} G(s,y)Q(\mathrm{d}s,\mathrm{d}y) = G(t,z) + \int_{[0,T]\times\mathbb{R}} D_{t,z}G(s,y)Q(\mathrm{d}s,\mathrm{d}y), \quad \mathbb{P}\text{-a.s.},$$

and $\int_{[0,T]\times\mathbb{R}} D_{t,z}G(s,y)Q(\mathrm{d}s,\mathrm{d}y)$ is a stochastic integral in Itô sense.

Proof. By square integrability of G, for q-a.e. $(s,y) \in [0,T]\times\mathbb{R}$, the chaotic decomposition property yields the unique representation $G(s,y) = \sum_{n=0}^{\infty} I_n((g_n(s,y),.))$, $g_n \in L^2_{T,q,n+1}$, $n \geq 0$. Square integrability and predictability of G implies that the stochastic integral $\int_{[0,T]\times\mathbb{R}} G(s,y)Q(\mathrm{d}s,\mathrm{d}y)$ is well-defined in the Itô sense and the Skorokhod integral, which coincides under the given assumptions with the Itô integral (see Theorem 6.1 in [24]) can be defined by the series expansion $\int_{[0,T]\times\mathbb{R}} G(s,y)Q(\mathrm{d}s,\mathrm{d}y) = \sum_{n=0}^{\infty} I_{n+1}(\hat{g}_n)$ according to Definition 3.1.4. The Skorokhod integral is Malliavin differentiable if and only if $\sum_{n=1}^{\infty}(n+1)(n+1)!\|\hat{g}_n\|^2_{L^2_{T,q,n+1}} < \infty$; see Definition 3.1.2. This series converges if and only if $G \in \mathbb{L}^{1,2}(\mathbb{R})$, by Definition 3.1.3.

Following Section 6 in [24], we can conclude that the required differentiation rule holds. To prove that the integral $\int_{[0,T]\times\mathbb{R}} D_{t,z}G(s,y)Q(\mathrm{d}s,\mathrm{d}y)$ is well-defined in the Itô sense, it is sufficient to show that the integrand $(\omega,s,y) \mapsto D_{t,z}G(s,y)(\omega)$ is a predictable mapping on $\Omega\times[0,T]\times\mathbb{R}$, as square integrability is already satisfied by $G \in \mathbb{L}^{1,2}(\mathbb{R})$. For q-a.e. $(s,y) \in [0,T]\times\mathbb{R}$, predictability of G implies that

$$G(s,y) = \sum_{n=0}^{\infty} I_n(g_n((s,y),.)) = \sum_{n=0}^{\infty} I_n(g_n((s,y),.)\mathbf{1}^{\otimes n}_{[0,s)}(.)), \quad \mathbb{P}\text{-a.s.},$$

and applying Definition 3.1.2 of the Malliavin derivative yields

$$D_{t,z}G(s,y) = \sum_{n=0}^{\infty} nI_{n-1}(g_n((s,y),(t,z),.)\mathbf{1}^{\otimes n}_{[0,s)}((t,z),.)),$$
$$\text{for } q\otimes q\text{-a.e. } ((t,z),(s,y)) \in ([0,T]\times\mathbb{R})^2,\ \mathbb{P}\text{-a.s.},$$

from which the required predictability of the integrand follows. As a by-product, let us note that $(\omega, s, y, t, z) \mapsto D_{t,z}G(s, y)(\omega)$ is jointly measurable. □

4. Variational differentiability of a solution

The main goal of this paper is to investigate Malliavin's differentiability of a solution of a backward stochastic differential equation with a time delayed generator. In this section, additionally to (A1)–(A5), we assume that

(A6) the generator f is of the following form

$$f(t, y_t, z_t, u_t) := f\Bigg(\omega, t, \int_{-T}^{0} y(t+v)\alpha(\mathrm{d}v), \int_{-T}^{0} z(t+v)\alpha(\mathrm{d}v) \int_{-T}^{0}\int_{\mathbb{R}-\{0\}} u(t+v, z)m(\mathrm{d}z)\alpha(\mathrm{d}v)\Bigg),$$

with a product measurable function $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, which is Lipschitz continuous in the last three variables for $\mathbb{P} \otimes \lambda$-a.e. $(\omega, t) \in \Omega \times [0, T]$, more precisely the generator satisfies (A3) with the same constant K,

(A7) the terminal value is Malliavin differentiable, i.e. $\xi \in \mathbb{D}^{1,2}(\mathbb{R})$, and

$$\mathbb{E}\left[\int_{[0,T]\times\mathbb{R}} |D_{s,z}\xi|^2 q(\mathrm{d}s, \mathrm{d}z)\right] < \infty,$$
$$\lim_{\epsilon\downarrow 0}\mathbb{E}\left[\int_0^T \int_{|z|\le\epsilon} |D_{s,z}\xi|^2 m(\mathrm{d}z)\mathrm{d}s\right] = 0,$$

(A8) for $\mathbb{P}\otimes\lambda$-a.e. $(\omega, t) \in \Omega \times [0, T]$, the mapping $(y, z, u) \mapsto f(\omega, t, y, z, u)$ is continuously differentiable in (y, z, u), with uniformly bounded and continuous partial derivatives f_y, f_z, f_u; we assume $f_y(\omega, t, ., ., .) = f_z(\omega, t, ., ., .) = f_u(\omega, t, ., ., .) = 0$ for $\omega \in \Omega, t < 0$;

(A9) for $(t, y, z, u) \in [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ we have $f(\cdot, t, y, z, u) \in \mathbb{D}^{1,2}(\mathbb{R})$ and

$$\mathbb{E}\left[\int_{[0,T]\times\mathbb{R}} \int_0^T \left|D_{s,z} f(\cdot, t, 0, 0, 0)\right|^2 \mathrm{d}t\, q(\mathrm{d}s, \mathrm{d}y)\right] < \infty,$$
$$\left|D_{s,z} f(\omega, t, \hat{y}, \hat{z}, \hat{u})\right| - \left|D_{s,z} f(\omega, t, \tilde{y}, \tilde{z}, \tilde{u})\right| \le L\left(|\hat{y} - \tilde{y}| + |\hat{z} - \tilde{z}| + |\hat{u} - \tilde{u}|\right),$$

$(s, z) \in [0, T] \times \mathbb{R}$, $(\hat{y}, \hat{z}, \hat{u}) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, $(\tilde{y}, \tilde{z}, \tilde{u}) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, for $\mathbb{P} \otimes \lambda$-a.e. $(\omega, t) \in \Omega \times [0, T]$.

The assumptions (A7)–(A9) are classical when dealing with Malliavin's differentiability; see Proposition 5.3 in [12] or Theorem 3.3.1 in [14] in the case of BSDEs driven by Brownian motions. We also remark that the generator in (A6) depends on $\int_{-T}^{0}\int_{\mathbb{R}-\{0\}} u(t+v, z)m(\mathrm{d}z)\alpha(\mathrm{d}v)$, which corresponds to a standard form of dependence appearing in BSDEs without delays and with jumps; see Proposition 2.6 and Remark 2.7 in [5].

We can state our main theorem.

Theorem 4.1. *Assume that* (A1)–(A9) *hold and that time horizon T and Lipschitz constant K of the generator f are sufficiently small, such that for some $\beta > 0$*

$$\delta := \delta(T, K, \beta, \alpha) = \left(8T + \frac{1}{\beta}\right) K \int_{-T}^{0} \mathrm{e}^{-\beta v}\alpha(\mathrm{d}v) \max\{1, T\} < 1.$$

1. *There exists a unique solution* $(Y, Z, U) \in \mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2_m(\mathbb{R})$ *of the time delayed BSDE*

$$\begin{aligned} Y(t) &= \xi + \int_t^T f\Bigg(\omega, r, \int_{-T}^0 Y(r+v)\alpha(\mathrm{d}v), \\ &\quad \int_{-T}^0 Z(r+v)\alpha(\mathrm{d}v), \int_{-T}^0 \int_{\mathbb{R}-\{0\}} U(r+v,z)m(\mathrm{d}z)\alpha(\mathrm{d}v)\Bigg)\mathrm{d}r \\ &\quad - \int_t^T Z(r)\mathrm{d}W(r) - \int_t^T \int_{\mathbb{R}-\{0\}} U(r,y)\tilde{M}(\mathrm{d}r,\mathrm{d}y), \quad 0 \le t \le T. \end{aligned} \tag{4.1}$$

2. *There exists a unique solution* $(Y^{s,0}, Z^{s,0}, U^{s,0}) \in \mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2_m(\mathbb{R})$ *of the time delayed BSDE*

$$\begin{aligned} Y^{s,0}(t) &= D_{s,0}\xi + \int_t^T f^{s,0}(r)\mathrm{d}r - \int_t^T Z^{s,0}(r)\mathrm{d}W(r) \\ &\quad - \int_t^T \int_{\mathbb{R}-\{0\}} U^{s,0}(r,y)\tilde{M}(\mathrm{d}r,\mathrm{d}y), \quad 0 \le s \le t \le T, \end{aligned} \tag{4.2}$$

with the generator

$$\begin{aligned} f^{s,0}(r) &= D_{t,0}f\Bigg(\omega, r, \int_{-T}^0 Y(r+v)\alpha(\mathrm{d}v), \int_{-T}^0 Z(r+v)\alpha(\mathrm{d}v), \\ &\quad \int_{-T}^0 \int_{\mathbb{R}-\{0\}} U(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg) + f_y\Bigg(\omega, r, \int_{-T}^0 Y(r+v)\alpha(\mathrm{d}v), \\ &\quad \int_{-T}^0 Z(r+v)\alpha(\mathrm{d}v), \int_{-T}^0 \int_{\mathbb{R}-\{0\}} U(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg) \\ &\quad \times \int_{-T}^0 Y^{s,0}(r+v)\alpha(\mathrm{d}v) + f_z\Bigg(\omega, r, \int_{-T}^0 Y(r+v)\alpha(\mathrm{d}v), \\ &\quad \int_{-T}^0 Z(r+v)\alpha(\mathrm{d}v), \int_{-T}^0 \int_{\mathbb{R}-\{0\}} U(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg) \\ &\quad \times \int_{-T}^0 Z^{s,0}(r+v)\alpha(\mathrm{d}v) + f_u\Bigg(\omega, r, \int_{-T}^0 Y(r+v)\alpha(\mathrm{d}v), \\ &\quad \int_{-T}^0 Z(r+v)\alpha(\mathrm{d}v), \int_{-T}^0 \int_{\mathbb{R}-\{0\}} U(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg) \\ &\quad \times \int_{-T}^0 \int_{\mathbb{R}-\{0\}} U^{s,0}(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v). \end{aligned} \tag{4.3}$$

3. *There exists a unique solution* $(Y^{s,z}, Z^{s,z}, U^{s,z}) \in \mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2_m(\mathbb{R})$ *of the time delayed BSDE*

$$Y^{s,z}(t) = D_{s,z}\xi + \int_t^T f^{s,z}(r)\mathrm{d}r - \int_t^T Z^{s,z}(r)\mathrm{d}W(r)$$

$$-\int_t^T \int_{\mathbb{R}-\{0\}} U^{s,z}(r,y)\tilde{M}(\mathrm{d}r,\mathrm{d}y), \quad 0 \le s \le t \le T, z \ne 0, \tag{4.4}$$

with the generator

$$\begin{aligned} f^{s,z}(r) = \Bigg\{ f\Bigg(&\omega^{s,z}, r, z\int_{-T}^0 Y^{s,z}(r+v)\alpha(\mathrm{d}v) + \int_{-T}^0 Y(r+v)\alpha(\mathrm{d}v),\\ &z\int_{-T}^0 Z^{s,z}(r+v)\alpha(\mathrm{d}v) + \int_{-T}^0 Z(r+v)\alpha(\mathrm{d}v),\\ &z\int_{-T}^0\int_{\mathbb{R}-\{0\}} U^{s,z}(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v)\\ &+\int_{-T}^0\int_{\mathbb{R}-\{0\}} U(r+v,x)m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg) - f\Bigg(\omega, r, \int_{-T}^0 Y(r+v)\alpha(\mathrm{d}v),\\ &\int_{-T}^0 Z(r+v)\alpha(\mathrm{d}v), \int_{-T}^0\int_{\mathbb{R}-\{0\}} U(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg)\Bigg\}\Big/ z, \end{aligned} \tag{4.5}$$

where we set

$$Y^{s,z}(t) = Z^{s,z}(t) = U^{s,z}(t,y) = 0, \quad (y,z) \in (\mathbb{R}-\{0\}) \times \mathbb{R}, \mathbb{P}\text{-a.s.}, \ t < s \le T. \tag{4.6}$$

Then $(Y, Z, U) \in \mathbb{L}^{1,2}(\mathbb{R}) \times \mathbb{L}^{1,2}(\mathbb{R}) \times \mathbb{L}^{1,2}(\mathbb{R})$ *and*
$(Y^{s,z}(t), Z^{s,z}(t), U^{s,z}(t,y))_{0\le s,t\le T,(y,z)\in(\mathbb{R}-\{0\})\times\mathbb{R}}$ *is a version of*
$(D_{s,z}Y(t), D_{s,z}Z(t), D_{s,z}U(t,y))_{0\le s,t\le T,(y,z)\in(\mathbb{R}-\{0\})\times\mathbb{R}}$.

We recall that $D_{t,0}f(r,.,.,.,.)$ appearing as the first term in (4.3) is the Malliavin derivative of f with respect to ω, whereas $\omega^{s,z}$ appearing in (4.5) is defined in (3.4).

Proof. We follow the idea of the proofs of Proposition 5.3 in [12], or Theorem 3.3.1 in [14]. Let us denote by C a finite constant which may change from line to line.

Step (1) Given $\beta > 0$, we prove existence of unique solutions of the Eqs. (4.1), (4.2) and (4.4) for a time horizon T and a Lipschitz constant K fulfilling $\delta(T, K, \beta, \alpha) < 1$.
The existence of a unique solution $(Y, Z, U) \in \mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2_m(\mathbb{R})$ of (4.1) follows from Theorem 2.1, since the assumptions (A1)–(A5) are satisfied. Under the additional assumptions (A7)–(A9), the time delayed BSDEs (4.2) and (4.4), with the generators (4.3) resp. (4.5), fulfill the conditions of Theorem 2.1. In particular the corresponding generators are Lipschitz continuous with the same Lipschitz constant K that the generator f possesses. It is easy to see that the generators (4.3) and (4.5) have the same Lipschitz constant K in the sense of (A3). Hence, $\delta(T, K, \beta, \alpha) < 1$ holds simultaneously for all BSDEs (4.1), (4.2) and (4.4) and we conclude that for q-a.e. $(s, z) \in [0, T] \times \mathbb{R}$ there exists a unique solution $(Y^{s,z}, Z^{s,z}, U^{s,z}) \in \mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2_m(\mathbb{R})$ of (4.2) or (4.4) satisfying (4.6).

Step (2) Consider a sequence $(Y^n, Z^n, U^n)_{n\in\mathbb{N}}$, constructed by Picard iteration scheme, which converges to (Y, Z, U). In this step we show that $(Y^n, Z^n, U^n) \in \mathbb{L}^{1,2}(\mathbb{R}) \times \mathbb{L}^{1,2}(\mathbb{R}) \times \mathbb{L}^{1,2}(\mathbb{R})$ implies $(Y^{n+1}, Z^{n+1}, U^{n+1}) \in \mathbb{L}^{1,2}(\mathbb{R}) \times \mathbb{L}^{1,2}(\mathbb{R}) \times \mathbb{L}^{1,2}(\mathbb{R})$, and that from $\mathbb{E}\left[\int_{[0,T]} \sup_{t\in[0,T]} |D_{s,z}Y^n(t)|^2 q(\mathrm{d}s,\mathrm{d}z)\right] < \infty$ we can as well deduce that $\mathbb{E}\left[\int_{[0,T]\times\mathbb{R}} \sup_{t\in[0,T]} |D_{s,z}Y^{n+1}(t)|^2 q(\mathrm{d}s,\mathrm{d}z)\right] < \infty$.

For that purpose, we study the iterations

$$Y^{n+1}(t) = \xi + \int_t^T f^n(r)\mathrm{d}r - \int_t^T Z^{n+1}(r)\mathrm{d}W(r) - \int_t^T \int_{\mathbb{R}-\{0\}} U^{n+1}(r,y)\tilde{M}(\mathrm{d}r,\mathrm{d}y), \quad 0 \le t \le T, \tag{4.7}$$

where we denote

$$f^n(r) = f\left(r, \int_{-T}^0 Y^n(r+v)\alpha(\mathrm{d}v), \int_{-T}^0 Z^n(r+v)\alpha(\mathrm{d}v), \int_{-T}^0 \int_{\mathbb{R}-\{0\}} U^n(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v)\right).$$

We first establish Malliavin's differentiability of $\int_t^T f^n(r)\mathrm{d}r$ by applying Lemma 3.2. Notice that $Y^n(t) \in \mathbb{D}^{1,2}(\mathbb{R})$, for λ-a.e. $t \in [-T,T]$. Similarly to (2.2), we can derive

$$\begin{aligned}\int_0^T \mathbb{E}\left[\int_{-T}^0 |Y^n(r+v)|^2\alpha(\mathrm{d}v)\right]\mathrm{d}r &= \mathbb{E}\left[\int_{-T}^0 \int_0^T |Y^n(r+v)|^2\mathrm{d}r\alpha(\mathrm{d}v)\right] \\ &= \mathbb{E}\left[\int_{-T}^0 \int_v^{T+v} |Y^n(w)|^2\mathrm{d}w\alpha(\mathrm{d}v)\right] \\ &\le T\mathbb{E}\left[\sup_{w\in[0,T]} |Y^n(w)|^2\right] < \infty\end{aligned}$$

together with

$$\begin{aligned}&\int_0^T \mathbb{E}\left[\int_{[0,T]\times\mathbb{R}} \int_{-T}^0 |D_{s,z}Y^n(r+v)|^2\alpha(\mathrm{d}v)q(\mathrm{d}s,\mathrm{d}z)\right]\mathrm{d}r \\ &\quad \le T\mathbb{E}\left[\int_{[0,T]\times\mathbb{R}} \sup_{w\in[0,T]} |D_{s,z}Y^n(w)|^2 q(\mathrm{d}s,\mathrm{d}z)\right] < \infty.\end{aligned}$$

This provides the assumptions of Lemma 3.2, and for λ-a.e. $r \in [0,T]$ we have $\int_{-T}^0 Y^n(r+v)\alpha(\mathrm{d}v) \in \mathbb{D}^{1,2}(\mathbb{R})$, and furthermore

$$D_{s,z}\int_{-T}^0 Y^n(r+v)\alpha(\mathrm{d}v) = \int_{-T}^0 D_{s,z}Y^n(r+v)\alpha(\mathrm{d}v), \quad \mathbb{P}\text{-a.s.},$$

for $q\otimes\lambda$-a.e. $(s,z,r) \in [0,T]\times\mathbb{R}\times[0,T]$. In an analogous way we derive

$$\begin{aligned}&D_{s,z}\int_{-T}^0 Z^n(r+v)\alpha(\mathrm{d}v) = \int_{-T}^0 D_{s,z}Z^n(r+v)\alpha(\mathrm{d}v), \\ &D_{s,z}\int_{-T}^0 \int_{\mathbb{R}-\{0\}} U^n(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v) = \int_{-T}^0 \int_{\mathbb{R}-\{0\}} D_{s,z}U^n(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v),\end{aligned}$$

holds $\mathbb{P}$-a.s. for $q\otimes\lambda$-a.e. $(s,z,r) \in [0,T]\times\mathbb{R}\times[0,T]$. We claim that for λ-a.e. $r \in [0,T]$ the random variable $f^n(r) \in \mathbb{D}^{1,2}(\mathbb{R})$ and for $q\otimes\lambda$-a.e. $(s,z,r) \in [0,T]\times\mathbb{R}\times[0,T]$ we have

$$
\begin{aligned}
D_{s,0}f^n(r) = {} & D_{t,0}f\Bigg(\cdot, r, \int_{-T}^{0} Y^n(r+v)\alpha(\mathrm{d}v), \int_{-T}^{0} Z^n(r+v)\alpha(\mathrm{d}v), \\
& \int_{-T}^{0}\int_{\mathbb{R}-\{0\}} U^n(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg) + f_y\Bigg(\cdot, r, \int_{-T}^{0} Y^n(r+v)\alpha(\mathrm{d}v), \\
& \int_{-T}^{0} Z^n(r+v)\alpha(\mathrm{d}v), \int_{-T}^{0}\int_{\mathbb{R}-\{0\}} U^n(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg) \\
& \times \int_{-T}^{0} D_{s,0}Y^n(r+v)\alpha(\mathrm{d}v) + f_z\Bigg(\cdot, r, \int_{-T}^{0} Y^n(r+v)\alpha(\mathrm{d}v), \\
& \int_{-T}^{0} Z^n(r+v)\alpha(\mathrm{d}v), \int_{-T}^{0}\int_{\mathbb{R}-\{0\}} U^n(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg) \\
& \times \int_{-T}^{0} D_{s,0}Z^n(r+v)\alpha(\mathrm{d}v) + f_u\Bigg(\cdot, r, \int_{-T}^{0} Y^n(r+v)\alpha(\mathrm{d}v), \\
& \int_{-T}^{0} Z^n(r+v)\alpha(\mathrm{d}v), \int_{-T}^{0}\int_{\mathbb{R}-\{0\}} U^n(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg) \\
& \times \int_{-T}^{0}\int_{\mathbb{R}-\{0\}} D_{s,0}U^n(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v), \qquad (4.8)
\end{aligned}
$$

and for $z \neq 0$

$$
\begin{aligned}
D_{s,z}f^n(r) = {} & \Bigg\{ f\Bigg(\cdot^{s,z}, r, z\int_{-T}^{0} D_{s,z}Y^n(r+v)\alpha(\mathrm{d}v) + \int_{-T}^{0} Y^n(r+v)\alpha(\mathrm{d}v), \\
& z\int_{-T}^{0} D_{s,z}Z^n(r+v)\alpha(\mathrm{d}v) + \int_{-T}^{0} Z^n(r+v)\alpha(\mathrm{d}v), \\
& z\int_{-T}^{0}\int_{\mathbb{R}-\{0\}} D_{s,z}U^n(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v) \\
& + \int_{-T}^{0}\int_{\mathbb{R}-\{0\}} U^n(r+v,x)m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg) - f\Bigg(\cdot, r, \int_{-T}^{0} Y^n(r+v)\alpha(\mathrm{d}v), \\
& \int_{-T}^{0} Z^n(r+v)\alpha(\mathrm{d}v), \int_{-T}^{0}\int_{\mathbb{R}-\{0\}} U^n(r+v,y)m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg)\Bigg\}\Bigg/ z. \qquad (4.9)
\end{aligned}
$$

The derivative (4.8) follows from the chain rule for the operator $D_{s,0}$, as for Theorem 2 in [21], whereas (4.9) follows from Proposition 5.5 in [24] provided that

$$
\begin{aligned}
& \mathbb{E}\left[\int_0^T |f^n(r)|^2 \mathrm{d}r\right] < \infty, \\
& \mathbb{E}\left[\int_0^T \int_0^T \int_{\mathbb{R}-\{0\}} |D_{s,z}f^n(r)|^2 m(\mathrm{d}z)\mathrm{d}s\mathrm{d}r\right] < \infty,
\end{aligned}
$$

hold. The finiteness of the first integral is obvious. The second integral can be shown to be finite by applying the Lipschitz continuity of the generator (A3), the Lipschitz continuity of the derivative of the function f with respect to ω and its square integrability (A9), as well as the assumption $(Y^n, Z^n, U^n) \in \mathbb{L}^{1,2}(\mathbb{R}) \times \mathbb{L}^{1,2}(\mathbb{R}) \times \mathbb{L}^{1,2}(\mathbb{R})$. Moreover,

$$\mathbb{E}\left[\int_0^T \int_0^T |D_{s,0} f^n(r)|^2 \mathrm{d}r \mathrm{d}s\right] < \infty,$$

and by Lemma 3.2 again we derive that for $0 \le t \le T$ we have $\xi + \int_t^T f^n(r)\mathrm{d}r \in \mathbb{D}^{1,2}(\mathbb{R})$ with Malliavin derivative

$$D_{s,z}\xi + \int_t^T D_{s,z} f^n(r)\mathrm{d}r, \quad q\text{-a.e. } (s,z) \in [0,T] \times \mathbb{R}, \tag{4.10}$$

where $D_{s,z} f^n$ is defined in (4.8) and (4.9). If we combine this result with Lemma 3.1, we can conclude

$$Y^{n+1}(t) = \mathbb{E}\left[\xi + \int_t^T f^n(r)\mathrm{d}r | \mathcal{F}_t\right] \in \mathbb{D}^{1,2}(\mathbb{R}), \quad 0 \le t \le T,$$

and from the Eq. (4.7) we derive

$$\int_t^T Z^{n+1}(r)\mathrm{d}W(r) \in \mathbb{D}^{1,2}(\mathbb{R}), \quad 0 \le t \le T, \tag{4.11}$$

and

$$\int_t^T \int_{\mathbb{R}-\{0\}} U^{n+1}(r,y)\tilde{M}(\mathrm{d}r, \mathrm{d}y) \in \mathbb{D}^{1,2}(\mathbb{R}), \quad 0 \le t \le T. \tag{4.12}$$

Therefore Lemma 3.3 yields $(Z^{n+1}, U^{n+1}) \in \mathbb{L}^{1,2}(\mathbb{R}) \times \mathbb{L}^{1,2}(\mathbb{R})$.
This allows us to differentiate the recursive equation (4.7) and obtain for q-a.e.$(s,z) \in [0,T]\times\mathbb{R}$

$$\begin{aligned} D_{s,z}Y^{n+1}(t) &= D_{s,z}\xi + \int_t^T D_{s,z} f^n(r)\mathrm{d}r - \int_t^T D_{s,z}Z^{n+1}(r)\mathrm{d}W(r) \\ &\quad - \int_t^T \int_{\mathbb{R}-\{0\}} D_{s,z}U^{n+1}(r,y)\tilde{M}(\mathrm{d}r,\mathrm{d}y), \quad s \le t \le T, \end{aligned} \tag{4.13}$$

and

$$D_{s,z}Y^{n+1}(t) = D_{s,z}Z^{n+1}(t) = D_{s,z}U^{n+1}(t,y) = 0, \quad t < s,\ y \in (\mathbb{R} - \{0\}). \tag{4.14}$$

Note that the time delayed BSDE (4.13) with generator (4.8) or (4.9) fulfills the assumptions of Theorem 2.1 with zero corresponding Lipschitz constant. We conclude that for q-a.e.$(s,z) \in [0,T] \times \mathbb{R}$ there exists a unique solution $(D_{s,z}Y^{n+1}, D_{s,z}Z^{n+1}, D_{s,z}U^{n+1}) \in \mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{H}^2_m$ of (4.13) satisfying (4.14). By applying Lemma 2.1, with $\tilde{\xi} = 0$ and $\tilde{f} = f$, together with the estimate (2.7), with $\delta = \delta(T, K, \beta, \alpha) < 1$ we derive the inequality

$$\begin{aligned} &\|D_{s,z}Y^{n+1}\|^2_{\mathbb{S}^2} + \|D_{s,z}Z^{n+1}\|^2_{\mathbb{H}^2} + \|D_{s,z}U^{n+1}\|^2_{\mathbb{H}^2_m} \\ &\quad \le 9\mathrm{e}^{\beta T}\mathbb{E}\left[|D_{s,z}\xi|^2\right] + \delta\left(\|D_{s,z}Y^n\|^2_{\mathbb{S}^2} + \|D_{s,z}Z^n\|^2_{\mathbb{H}^2} + \|D_{s,z}U^n\|^2_{\mathbb{H}^2_m}\right). \end{aligned} \tag{4.15}$$

This in turn yields $\mathbb{E}\left[\int_{[0,T]\times\mathbb{R}} \sup_{t\in[0,T]} |D_{s,z}Y^{n+1}(t)|^2 q(\mathrm{d}s,\mathrm{d}z)\right] < \infty$, and in particular, $Y^{n+1} \in \mathbb{L}^{1,2}(\mathbb{R})$.

Step (3) We establish the integrability of the solution $Y^{s,z}(t)$, $Z^{s,z}(t)$, $U^{s,z}(t,y)$ with respect to the product measure q on $([0,T]\times\mathbb{R})^2$.
Take $(s,z)\in[0,T]\times\mathbb{R}$. Consider the unique solution $(Y^{s,z},Z^{s,z},U^{s,z})\in\mathbb{S}^2(\mathbb{R})\times\mathbb{H}^2(\mathbb{R})\times\mathbb{H}^2_m(\mathbb{R})$ of the Eq. (4.2) or (4.4). Lemma 2.1, with $\tilde{\xi}=0$ and $\tilde{f}=f$ together with the estimates (2.7) and (2.8) yield the inequality

$$\begin{aligned}&\|Y^{s,z}\|^2_{\mathbb{S}^2}+\|Z^{s,z}\|^2_{\mathbb{H}^2}+\|U^{s,z}\|^2_{\mathbb{H}^2_m}\\&\quad\le 9\mathrm{e}^{\beta T}\mathbb{E}\left[|D_{s,z}\xi|^2\right]+\delta\left(\|Y^{s,z}\|^2_{\mathbb{S}^2}+\|Z^{s,z}\|^2_{\mathbb{H}^2}+\|U^{s,z}\|^2_{\mathbb{H}^2_m}\right),\end{aligned}$$

so that under $\delta=\delta(T,K,\beta,\alpha)<1$ we obtain for q-a.e. $(s,z)\in[0,T]\times\mathbb{R}$

$$\|Y^{s,z}\|^2_{\mathbb{S}^2}+\|Z^{s,z}\|^2_{\mathbb{H}^2}+\|U^{s,z}\|^2_{\mathbb{H}^2_m}\le C\mathbb{E}\left[|D_{s,z}\xi|^2\right], \tag{4.16}$$

and we arrive at

$$\begin{aligned}&\mathbb{E}\left[\int_{([0,T]\times\mathbb{R})^2}|Y^{s,z}(t)|^2q(\mathrm{d}t,\mathrm{d}y)q(\mathrm{d}s,\mathrm{d}z)\right]<\infty,\\&\mathbb{E}\left[\int_{([0,T]\times\mathbb{R})^2}|Z^{s,z}(t)|^2q(\mathrm{d}t,\mathrm{d}y)q(\mathrm{d}s,\mathrm{d}z)\right]<\infty,\\&\mathbb{E}\left[\int_{([0,T]\times\mathbb{R})^2}|U^{s,z}(t,y)|^2q(\mathrm{d}t,\mathrm{d}y)q(\mathrm{d}s,\mathrm{d}z)\right]<\infty.\end{aligned}$$

Step (4) We show convergence of $(Y^n,Z^n,U^n)_{n\in\mathbb{N}}$ in $\mathbb{L}^{1,2}(\mathbb{R})\times\mathbb{L}^{1,2}(\mathbb{R})\times\mathbb{L}^{1,2}(\mathbb{R})$.
From Theorem 2.1 we already know that $(Y^n,Z^n,U^n)_{n\in\mathbb{N}}$ converges in $\mathbb{S}^2(\mathbb{R})\times\mathbb{H}^2(\mathbb{R})\times\mathbb{H}^2_m(\mathbb{R})$. We have to prove that the corresponding Malliavin derivatives converge. The convergence

$$\begin{aligned}&\lim_{n\to\infty}\int_{[0,T]\times\mathbb{R}}\Big(\|Y^{s,z}-D_{s,z}Y^{n+1}\|^2_{\mathbb{H}^2}+\|Z^{s,z}-D_{s,z}Z^{n+1}\|^2_{\mathbb{H}^2}\\&\quad+\ \|U^{s,z}-D_{s,z}U^{n+1}\|^2_{\mathbb{H}^2_m}\Big)q(\mathrm{d}s,\mathrm{d}z)=0,\end{aligned}$$

for $z=0$ can be proved in the similar way as in the case of a BSDE without delay driven by a Brownian motion; see for example Theorem 3.3.1 in [14]. We only prove the convergence for $z\neq 0$.
Lemma 2.1, applied to the time delayed BSDEs (4.4) and (4.13) with (4.14), yield the inequality

$$\begin{aligned}&\|Y^{s,z}-D_{s,z}Y^{n+1}\|^2_{\mathbb{S}^2}+\|Z^{s,z}-D_{s,z}Z^{n+1}\|^2_{\mathbb{H}^2}+\|U^{s,z}-D_{s,z}U^{n+1}\|^2_{\mathbb{H}^2_m}\\&\quad\le\left(8T+\frac{1}{\beta}\right)\mathbb{E}\left[\int_s^T\mathrm{e}^{\beta r}|f^{s,z}(r)-D_{s,z}f^n(r)|^2\mathrm{d}r\right],\end{aligned} \tag{4.17}$$

for q-a.e. $(s,z)\in[0,T]\times\mathbb{R}$.
First, by the Lipschitz continuity condition (A3) for the generator f and the Lipschitz continuity condition (A9) for the derivative of the function f with respect to ω we obtain for $\lambda\otimes m\otimes\lambda$-a.e. $(s,z,r)\in[0,T]\times(\mathbb{R}-\{0\})\times[0,T]$ the following two estimates

$$\begin{aligned}|f^{s,z}(r)-D_{s,z}f^n(r)|^2\le C\Bigg(&\int_{-T}^0|Y^{s,z}(r+v)|^2\alpha(\mathrm{d}v)+\int_{-T}^0|Z^{s,z}(r+v)|^2\alpha(\mathrm{d}v)\\&+\int_{-T}^0\int_{\mathbb{R}-\{0\}}|U^{s,z}(r+v,y)|^2m(\mathrm{d}y)\alpha(\mathrm{d}v)\end{aligned}$$

$$+ \int_{-T}^{0} |D_{s,z}Y^n(r+v)|^2\alpha(\mathrm{d}v) + \int_{-T}^{0} |D_{s,z}Z^n(r+v)|^2\alpha(\mathrm{d}v)$$
$$+ \int_{-T}^{0}\int_{\mathbb{R}-\{0\}} |D_{s,z}U^n(r+v,y)|^2 m(\mathrm{d}y)\alpha(\mathrm{d}v)$$
$$+ \int_{-T}^{0} |Y^n(r+v) - Y(r+v)|^2\alpha(\mathrm{d}v) + \int_{-T}^{0} |Z^n(r+v) - Z(r+v)|^2\alpha(\mathrm{d}v)$$
$$+ \int_{-T}^{0}\int_{\mathbb{R}-\{0\}} |U^n(r+v,y) - U(r+v,y)|^2 m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg), \tag{4.18}$$

and for any $\lambda > 0$

$$|f^{s,z}(r) - D_{s,z}f^n(r)|^2 \le \left(1+\frac{1}{\lambda}\right)^2 K\Bigg(\int_{-T}^{0} |Y^{s,z}(r+v) - D_{s,z}Y^n(r+v)|^2\alpha(\mathrm{d}v)$$
$$+ \int_{-T}^{0} |Z^{s,z}(r+v) - D_{s,z}Z^n(r+v)|^2\alpha(\mathrm{d}v)$$
$$+ \int_{-T}^{0}\int_{\mathbb{R}-\{0\}} |U^{s,z}(r+v,y) - D_{s,z}U^n(r+v,y)|^2 m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg)$$
$$+ (1+\lambda)\left(2+\frac{1}{\lambda}\right) K\Bigg(\int_{-T}^{0} |Y(r+v) - Y^n(r+v)|^2\alpha(\mathrm{d}v)$$
$$+ \int_{-T}^{0} |Z(r+v) - Z^n(r+v)|^2\alpha(\mathrm{d}v)$$
$$+ \int_{-T}^{0}\int_{\mathbb{R}-\{0\}} |U(r+v,y) - U^n(r+v,y)|^2 m(\mathrm{d}y)\alpha(\mathrm{d}v)\Bigg)\Big/ z^2. \tag{4.19}$$

Note that

$$\int_{[0,T]\times(\mathbb{R}-\{0\})}\Bigg(\|Y^{s,z} - D_{s,z}Y^{n+1}\|^2_{\mathbb{S}^2}$$
$$+ \|Z^{s,z} - D_{s,z}Z^{n+1}\|^2_{\mathbb{H}^2} + \|U^{s,z} - D_{s,z}U^{n+1}\|^2_{\mathbb{H}^2_m}\Bigg) q(\mathrm{d}s,\mathrm{d}z)$$
$$= \lim_{\epsilon\downarrow 0}\int_0^T\int_{|z|>\epsilon}\Bigg(\|Y^{s,z} - D_{s,z}Y^{n+1}\|^2_{\mathbb{H}^2}$$
$$+ \|Z^{s,z} - D_{s,z}Z^{n+1}\|^2_{\mathbb{H}^2} + \|U^{s,z} - D_{s,z}U^{n+1}\|^2_{\mathbb{H}^2_m}\Bigg) m(\mathrm{d}z)\mathrm{d}s. \tag{4.20}$$

We prove that this convergence is uniform in n.
Choose $\varepsilon > 0$ sufficiently small. By assumption (A7) we can find $\bar{\epsilon}$ such that

$$\mathbb{E}\left[\int_0^T\int_{|z|\le\bar{\epsilon}} |D_{s,z}\xi|^2 m(\mathrm{d}z)\mathrm{d}s\right] < \varepsilon,$$

and

$$\int_{|z|\le\bar{\epsilon}} m(\mathrm{d}z) < \varepsilon.$$

Take arbitrary $0 < \epsilon_1 < \epsilon_2 \le \bar{\epsilon}$. By applying the inequality (4.17), the estimate (4.18) and by similar calculations as in (2.7) we can derive

$$\begin{aligned}
&\int_0^T \int_{\epsilon_1<|z|\le\epsilon_2} \Big(\|Y^{s,z} - D_{s,z}Y^{n+1}\|^2_{\mathbb{S}^2} + \|Z^{s,z} - D_{s,z}Z^{n+1}\|^2_{\mathbb{H}^2} \\
&\quad + \|U^{s,z} - D_{s,z}U^{n+1}\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s \\
&\le C\int_0^T \int_{\epsilon_1<|z|\le\epsilon_2} \mathbb{E}\left[\int_s^T \mathrm{e}^{\beta r}|f^{s,z}(r) - D_{s,z}f^n(r)|^2\mathrm{d}r\right] m(\mathrm{d}z)\mathrm{d}s \\
&\le C\Bigg\{\int_0^T \int_{\epsilon_1<|z|\le\epsilon_2} \Big(\|Y^{s,z}\|^2_{\mathbb{S}^2} + \|Z^{s,z}\|^2_{\mathbb{H}^2} + \|U^{s,z}\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s \\
&\quad + \int_0^T \int_{\epsilon_1<|z|\le\epsilon_2} \Big(\|D_{s,z}Y^n\|^2_{\mathbb{S}^2} + \|D_{s,z}Z^n\|^2_{\mathbb{H}^2} + \|D_{s,z}U^n\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s \\
&\quad + \int_0^T \int_{\epsilon_1<|z|\le\epsilon_2} \Big(\|Y^n - Y\|^2_{\mathbb{S}^2} + \|Z^n - Z\|^2_{\mathbb{H}^2} + \|U^n - U\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s\Bigg\}. \qquad (4.21)
\end{aligned}$$

To estimate the first term on the right hand side of (4.21), notice that the inequality (4.16) yields

$$\begin{aligned}
&\int_0^T \int_{\epsilon_1<|z|\le\epsilon_2} \Big(\|Y^{s,z}\|^2_{\mathbb{S}^2} + \|Z^{s,z}\|^2_{\mathbb{H}^2} + \|U^{s,z}\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s \\
&\le C\mathbb{E}\left[\int_0^T \int_{\epsilon_1<|z|\le\epsilon_2} |D_{s,z}\xi|^2 m(\mathrm{d}z)\mathrm{d}s\right] < C\varepsilon. \qquad (4.22)
\end{aligned}$$

Recalling $\delta = \delta(T, K, \beta, \alpha) < 1$ and applying the inequality (4.15) we estimate the second term in (4.21) by

$$\begin{aligned}
&\int_0^T \int_{\epsilon_1<|z|\le\epsilon_2} \Big(\|D_{s,z}Y^n\|^2_{\mathbb{S}^2} + \|D_{s,z}Z^n\|^2_{\mathbb{H}^2} + \|D_{s,z}U^n\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s \\
&\le 9\mathrm{e}^{\beta T}\mathbb{E}\left[\int_0^T \int_{\epsilon_1<|z|\le\epsilon_2} |D_{s,z}\xi|^2 m(\mathrm{d}z)\mathrm{d}s\right] \\
&\quad + \delta\int_0^T \int_{\epsilon_1<|z|\le\epsilon_2} \Big(\|D_{s,z}Y^{n-1}\|^2_{\mathbb{S}^2} + \|D_{s,z}Z^{n-1}\|^2_{\mathbb{H}^2} + \|D_{s,z}U^{n-1}\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s \\
&< \frac{9\mathrm{e}^{\beta T}\varepsilon}{1-\delta} + \delta^n\int_0^T \int_{\epsilon_1<|z|\le\epsilon_2} \Big(\|D_{s,z}Y^0\|^2_{\mathbb{S}^2} + \|D_{s,z}Z^0\|^2_{\mathbb{H}^2} + \|D_{s,z}U^0\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s. \\
&\qquad (4.23)
\end{aligned}$$

The estimate of the third term follows from the contraction inequality (2.8)

$$\begin{aligned}
&\int_0^T \int_{\epsilon_1<|z|\le\epsilon_2} \Big(\|Y^n - Y\|^2_{\mathbb{S}^2} + \|Z^n - Z\|^2_{\mathbb{H}^2} + \|U^n - U\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s \\
&\le \delta^n T\Big(\|Y^0 - Y\|^2_{\mathbb{S}^2} + \|Z^0 - Z\|^2_{\mathbb{H}^2} + \|U^0 - U\|^2_{\mathbb{H}^2_m}\Big)\varepsilon. \qquad (4.24)
\end{aligned}$$

Choosing $Y^0 = Z^0 = U^0 = 0$ and combining (4.22)–(4.24) gives the uniform convergence of (4.20).

Next, by applying the inequality (4.17), the estimate (4.19) and similar calculations as in (2.7) and (2.8) we can derive

$$\begin{aligned}
&\int_0^T \int_{|z|>\epsilon} \Big(\|Y^{s,z} - D_{s,z}Y^{n+1}\|^2_{\mathbb{S}^2} + \|Z^{s,z} - D_{s,z}Z^{n+1}\|^2_{\mathbb{H}^2} \\
&\quad + \|U^{s,z} - D_{s,z}U^{n+1}\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s \\
&\le \left(8T + \frac{1}{\beta}\right) \int_0^T \int_{|z|>\epsilon} \mathbb{E}\left[\int_s^T \mathrm{e}^{\beta r} |f^{s,z}(r) - D_{s,z}f^n(r)|^2 \mathrm{d}r\right] m(\mathrm{d}z)\mathrm{d}s \\
&\le \delta \Bigg\{ \left(1 + \frac{1}{\lambda}\right)^2 \int_0^T \int_{|z|>\epsilon} \Bigg(\|Y^{s,z} - D_{t,z}Y^n\|^2_{\mathbb{S}^2} \\
&\quad + \|Z^{s,z} - D_{s,z}Z^n\|^2_{\mathbb{H}^2} + \|U^{s,z} - D_{s,z}U^n\|^2_{\mathbb{H}^2_m}\Bigg) m(\mathrm{d}z)\mathrm{d}s \\
&\quad + (1+\lambda)\left(2 + \frac{1}{\lambda}\right)\left(\|Y^n - Y\|^2_{\mathbb{S}^2} + \|Z^n - Z\|^2_{\mathbb{H}^2} + \|U^n - U\|^2_{\mathbb{H}^2_m}\right) \int_{|z|>\epsilon} \nu(\mathrm{d}z)\Bigg\},
\end{aligned}$$

and we choose λ sufficiently large such that $\tilde{\delta} := \delta\left(1 + \frac{1}{\lambda}\right)^2 < 1$.

Due to the convergence of $(Y^n, Z^n, U^n)_{n\in\mathbb{N}}$, for a sufficiently small $\varepsilon > 0$ we can find N sufficiently large such that for all $n \ge N$

$$(1+\lambda)\left(2 + \frac{1}{\lambda}\right)\left(\|Y^n - Y\|^2_{\mathbb{S}^2} + \|Z^n - Z\|^2_{\mathbb{H}^2} + \|U^n - U\|^2_{\mathbb{H}^2_m}\right) \int_{|z|>\epsilon} \nu(\mathrm{d}z) < \varepsilon.$$

We derive the recursion for $n \ge N$

$$\begin{aligned}
&\int_0^T \int_{|z|>\epsilon} \Big(\|Y^{s,z} - D_{s,z}Y^{n+1}\|^2_{\mathbb{S}^2} + \|Z^{s,z} - D_{s,z}Z^{n+1}\|^2_{\mathbb{H}^2} \\
&\quad + \|U^{s,z} - D_{s,z}U^{n+1}\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s \\
&< \tilde{\delta} \Bigg\{ \int_0^T \int_{|z|>\epsilon} \Big(\|Y^{s,z} - D_{s,z}Y^n\|^2_{\mathbb{S}^2} + \|Z^{s,z} - D_{s,z}Z^n\|^2_{\mathbb{H}^2} \\
&\quad + \|U^{s,z} - D_{s,z}U^n\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s \Bigg\} + \delta\varepsilon \\
&< \tilde{\delta}^{n-N} \int_0^T \int_{|z|>\epsilon} \Big(\|Y^{s,z} - D_{s,z}Y^N\|^2_{\mathbb{S}^2} + \|Z^{s,z} - D_{s,z}Z^N\|^2_{\mathbb{H}^2} \\
&\quad + \|U^{s,z} - D_{s,z}U^N\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s + \frac{\delta\varepsilon}{1 - \tilde{\delta}},
\end{aligned}$$

and finally we conclude that

$$\begin{aligned}
&\lim_{n\to\infty} \int_0^T \int_{|z|>\epsilon} \Big(\|Y^{s,z} - D_{s,z}Y^{n+1}\|^2_{\mathbb{S}^2} + \|Z^{s,z} - D_{s,z}Z^{n+1}\|^2_{\mathbb{H}^2} \\
&\quad + \|U^{s,z} - D_{s,z}U^{n+1}\|^2_{\mathbb{H}^2_m}\Big) m(\mathrm{d}z)\mathrm{d}s = 0.
\end{aligned}$$

The equation

$$\lim_{n\to\infty}\int_{[0,T]\times(\mathbb{R}-\{0\})}\Big(\|Y^{s,z}-D_{s,z}Y^{n+1}\|^2_{\mathbb{S}^2}+\|Z^{s,z}-D_{s,z}Z^{n+1}\|^2_{\mathbb{H}^2}$$
$$+\ \|U^{s,z}-D_{s,z}U^{n+1}\|^2_{\mathbb{H}^2_m}\Big)q(\mathrm{d}s,\mathrm{d}z)=0$$

now follows by interchanging the limits in n and ε in (4.20).

Step (4) Since the space $\mathbb{L}^{1,2}(\mathbb{R})$ is a Hilbert space and the Malliavin derivative is a closed operator, see Theorem 12.6 in [11], the claim that $(Y,Z,U)\in\mathbb{L}^{1,2}(\mathbb{R})\times\mathbb{L}^{1,2}(\mathbb{R})\times\mathbb{L}^{1,2}(\mathbb{R})$ and $(Y^{s,z}(t),Z^{s,z}(t),U^{s,z}(t,y))_{0\le s,t\le T,(y,z)\in(\mathbb{R}-\{0\})\mathbb{R}}$ is a version of the derivative $(D_{s,z}Y(t),D_{s,z}Z(t),D_{s,z}U(t,y))_{0\le s,t\le T,(y,z)\in(\mathbb{R}-\{0\})\mathbb{R}}$ follows, and finishes the proof. □

The following corollary shows that the interpretation of the solution component (Z,U) in terms of the Malliavin trace of Y still holds for BSDEs with time delayed generators.

Corollary 4.1. *Under the assumptions of Theorem 4.1, we have*

$$\Big((D_{t,0}Y)^{\mathcal{P}}(t)\Big)_{0\le t\le T}\ \textit{is a version of}\ \ (Z(t))_{0\le t\le T}\ ,$$
$$\Big((D_{t,z}Y)^{\mathcal{P}}(t)\Big)_{0\le t\le T,z\in(\mathbb{R}-\{0\})}\ \textit{is a version of}\ \ (U(t,z))_{0\le t\le T,z\in(\mathbb{R}-\{0\})}\ ,$$

where $(\cdot)^{\mathcal{P}}$ denotes the predictable projection of a process.

Proof. The solution of (4.1) satisfies

$$\begin{aligned}Y(s)&=Y(0)-\int_0^s f\Bigg(r,\int_{-T}^0 Y(r+v)\alpha(\mathrm{d}v),\\&\qquad\int_{-T}^0 Z(r+v)\alpha(\mathrm{d}v),\int_{-T}^0\int_{\mathbb{R}-\{0\}}U(r+v,y)m(\mathrm{d}y)\mathrm{d}v\Bigg)\mathrm{d}r\\&\qquad+\int_0^s Z(r)\mathrm{d}W(r)+\int_0^s\int_{\mathbb{R}-\{0\}}U(r,y)\tilde{M}(\mathrm{d}r,\mathrm{d}y),\quad 0\le s\le T.\end{aligned}\tag{4.25}$$

By differentiating (4.25) we obtain according to Lemma 3.3 for q-a.e. $(u,z)\in[0,T]\times\mathbb{R}$

$$\begin{aligned}D_{u,0}Y(s)&=Z(u)-\int_u^s D_{u,0}f(r)\mathrm{d}r+\int_u^s D_{u,0}Z(r)\mathrm{d}W(r)\\&\qquad+\int_u^s\int_{\mathbb{R}-\{0\}}D_{u,0}U(r,y)\tilde{M}(\mathrm{d}r,\mathrm{d}y),\quad 0\le u\le s\le T,\end{aligned}$$

and for $z\neq 0$

$$\begin{aligned}D_{u,z}Y(s)&=U(u,z)-\int_u^s D_{u,z}f(r)\mathrm{d}r+\int_u^s D_{u,z}Z(r)\mathrm{d}W(r)\\&\qquad+\int_u^s\int_{\mathbb{R}-\{0\}}D_{u,z}U(r,y)\tilde{M}(\mathrm{d}r,\mathrm{d}y),\quad 0\le u\le s\le T,\end{aligned}$$

where the derivative operators $D_{u,z}$ are defined by (4.3) and (4.5). Since the mappings $s\mapsto\int_u^s D_{u,z}f(r)\mathrm{d}r$, $s\mapsto\int_u^s D_{u,z}Z(r)\mathrm{d}W(r)$ are $\mathbb{P}$-a.s. continuous and the mapping $s\mapsto\int_u^s\int_{\mathbb{R}-\{0\}}$

$D_{u,z}U(r, y)\tilde{M}(\mathrm{d}r, \mathrm{d}y)$ is $\mathbb{P}$-a.s. càdlàg (see Theorems 4.2.12 and 4.2.14 in [4]), taking the limit $s \downarrow u$ yields

$$D_{u,0}Y(u) = Z(u), \quad \text{for } \lambda\text{-a.e. } u \in [0, T], \ \mathbb{P}\text{-a.s.,}$$
$$D_{u,z}Y(u) = U(u, z) \quad \text{for } \lambda \otimes m\text{-a.e. } (u, z) \in [0, T] \times (\mathbb{R} - \{0\}), \ \mathbb{P}\text{-a.s.}$$

As $Y \in \mathbb{S}^2(\mathbb{R})$ has $\mathbb{P}$-a.s. càdlàg $\mathbb{F}$-adapted trajectories, for $0 \leq u \leq T$ we have the representation

$$Y(u) = \sum_{n=0}^{\infty} I_n(g_n((u, 0), .)) = \sum_{n=0}^{\infty} I_n(g_n((u, 0), .)\mathbf{1}_{[0,u]}^{\otimes n}(.)), \quad g_n \in L^2_{T,q,n+1}, n \geq 0,$$

with càdlàg mappings $u \mapsto g_n((u, 0), .)$. By Definition 3.1.2 of the Malliavin derivative we arrive at

$$D_{u,z}Y(u) = \sum_{n=0}^{\infty} n I_{n-1}(g_n((u, 0), (u, z), .)\mathbf{1}_{[0,u]}^{\otimes n}((u, z), .)),$$
$$\text{for } q\text{-a.e. } (u, z) \in [0, T] \times \mathbb{R}.$$

For $\delta_{\{0\}} \times m$-a.e. $z \in \mathbb{R}$, we conclude that the mapping $(u, \omega) \mapsto D_{u,z}Y(u)(\omega)$ is $\mathbb{F}$-adapted and measurable and has a progressively measurable (optional) modification. Moreover, notice that the optional process $u \mapsto D_{u,z}Y(u)$ and its unique predictable projection $u \mapsto (D_{u,z}Y)^{\mathcal{P}}(u)$ are modifications of each other; see Theorem 5.5 in [13]. Finally, we remark that there exists a $\mathcal{P} \times \mathcal{B}(\mathbb{R})$ measurable version of $(\omega, u, z) \mapsto (D_{u,z}Y)^{\mathcal{P}}(u)(\omega)$; see Lemma 2.2 in [1]. This completes the proof. □

Acknowledgements

This paper was written while the first author was staying at Humboldt University, Berlin. Łukasz Delong acknowledges the financial support from the AMaMeF programme.

References

[1] S. Ankirchner, P. Imkeller, Quadratic hedging of weather and catastrophe risk by using short term climate predictions, HU Berlin, Preprint, 2008.
[2] S. Ankirchner, P. Imkeller, G. Reis, Classical and variational differentiability of BSDEs with quadratic growth, Electron. J. Probab. 12 (2007) 1418–1453.
[3] S. Ankirchner, P. Imkeller, G. Reis, Pricing and hedging of derivatives based on non-tradable underlyings, Math. Finance 20 (2010) 289–312.
[4] D. Applebaum, Lévy Processes and Stochastic Calculus, Cambridge University Press, Cambridge, 2004.
[5] G. Barles, R. Buckdahn, E. Pardoux, Backward stochastic differential eqautions and integral-partial differential equations, Stoch. Stoch. Rep. 60 (1997) 57–83.
[6] D. Becherer, Bounded solutions to backward SDE's with jumps for utility optimization and indifference pricing, Ann. Appl. Probab. 16 (4) (2006) 2027–2054.
[7] B. Bouchard, R. Elie, Discrete time approximation of decoupled forward–backward SDE with jumps, Stochastic Process. Appl. 118 (2008) 53–75.
[8] R. Buckdahn, P. Imkeller, Backward stochastic differential equations with time delayed generator, Preprint, 2008.
[9] Ł. Delong, Applications of time-delayed BSDEs to pricing, hedging and management of insurance and financial risks, Working Paper, 2010.
[10] Ł. Delong, P. Imkeller, Backward stochastic differential equations with time delayed generators—new results and counterexamples, Ann. Appl. Probab. (2010) (in press).
[11] G. Di Nunno, B. Øksendal, F. Proske, Malliavin Calculus for Lévy Processes with Applications to Finance, Springer-Verlag, 2009.

[12] N. El Karoui, S. Peng, M.C. Quenez, Backward stochastic differential equations in finance, Math. Finance 7 (1) (1997) 1–71.
[13] S. He, J. Wang, J. Yan, Semimartingale Theory and Stochastic Calculus, CRC Press Inc., 1992.
[14] P. Imkeller, Malliavin's Calculus and Applications in Stochastic Control and Finance, IM PAN Lectures Notes, Warsaw, 2008.
[15] P. Imkeller, G. Reis, Path regularity and explicit truncations order for BSDE with drivers of quadratic growth, HU Berlin, Preprint, 2009.
[16] I. Karatzas, D.L. Ocone, A generalized clark representation formula with application to optimal portfolios, Stochastics 34 (1992) 187–220.
[17] A. Kebaier, A. Kohatsu-Higa, An optimal control variance reduction method for density estimation, Stochastic Process. Appl. 118 (2008) 2143–2180.
[18] P. Malliavin, A. Thalmaier, Stochastic Calculus of Variations in Mathematical Finance, Springer-Verlag, Berlin, 2006.
[19] D. Nualart, The Malliavin Calculus and Related Topics, Springer-Verlag, 1995.
[20] E. Pardoux, S. Peng, Adapted solution of a backward stochastic differential equation, Syst. Control Lett. 14 (1990) 55–61.
[21] E. Petrou, Malliavin calculus in Lévy spaces and applications to finance, Electron. J. Probab. 13 (2008) 852–879.
[22] J. Picard, On the existence of smooth densities for jump processes, Probab. Theory Related Fields 105 (1996) 481–511.
[23] M. Royer, Backward stochastic differential equations with jumps and related non-linear expectations, Stochastic Process. Appl. 116 (2006) 1358–1376.
[24] J.L. Solé, F. Utzet, J. Vives, Canonical Lévy process and Malliavin calculus, Stochastic Process. Appl. 117 (2007) 165–187.